Schätzung von Kovarianzmatrizen in linearen Modellen

Dissertation
zur
Erlangung des Grades
eines Doktors der Naturwissenschaften
der Abteilung für Mathematik
der Ruhr-Universität Bochum

von
Wolfgang Urfer
aus
Stuttgart

Bochum 1975

Schätzung von Kovarianzmatrizen in linearen Modellen

Dissertation
zur
Erlangung des Grades
eines Doktors der Naturwissenschaften
der Abteilung für Mathematik
der Ruhr-Universität Bochum

von
Wolfgang Urfer
aus
Stuttgart

Bochum 1975

UNIV.
LEIDEN
BIBL.

Dissertation eingereicht am 15.10.1975
Tag der mündlichen Prüfung: 9.12.1975

Dekan: Prof. Dr. H. Amann
Referent: Prof. Dr. V. Baumann
Korreferent: Prof. Dr. P. Gänßler

Inhaltsverzeichnis

		Seite
	Einleitung	1
1.	Orthogonalprojektoren und die Darstellung von klassischen Schätzmethoden für Varianzkomponenten	
1.1.	Modell und Problemstellung	4
1.2.	Hilfsmittel, Orthogonalprojektoren und Erwartungswerte quadratischer Formen	6
1.3.	HENDERSON's Methode I im reinen Varianzkomponentenmodell	10
1.4.	HENDERSON's Methode III im gemischten Modell	14
2.	Neuere Methoden zur Schätzung von Kovarianzmatrizen und ihr Zusammenhang	
2.1.	MINQUE-Theorie zur Schätzung von Kovarianzmatrizen	18
2.2.	SEELY's Methode und ihr Zusammenhang mit der MINQUE-Theorie	22
2.3.	Weitere Zusammenhänge neuerer Schätzmethoden	30
2.4.	Untersuchungen bei hierarchischen Designs	37
2.5.	Schätzer von Varianzkomponenten und ihr GAUSS'sches Risiko	44
3.	Anwendung der Schätztheorie von Kovarianzmatrizen auf RCR-Modelle von P.A.V.B. SWAMY.	58
	Literaturverzeichnis	62

Einleitung

Die vorliegende Arbeit geht von einem allgemeinen univariaten linearen Modell aus, in dem sowohl die den Erwartungswert bestimmenden Parameter (feste Effekte) als auch unbekannte Varianz- und Kovarianzkomponenten vorkommen. Das Ziel der Untersuchung ist die Schätzung der Kovarianzmatrix eines Zufallsvektors im allgemeinen linearen Modell. Einerseits ist die damit erhaltene Information über die unbekannten Parameter der Kovarianzmatrix von Interesse, und andererseits ermöglicht eine Schätzung der Kovarianzmatrix die Berechnung eines AITKEN-Schätzers für den unbekannten Parametervektor der festen Effekte.

Nach der Zusammenstellung von Bezeichnungen und Hilfsmitteln aus der Theorie der endlich dimensionalen Vektorräume wird ein Satz über die Matrizendarstellung eines Orthogonalprojektors angegeben. Die für ein Beispiel eines reinen Varianzkomponentenmodells von S.R. SEARLE in [9] angegebene Methode I von HENDERSON wird geometrisch dargestellt, und es wird gezeigt, daß diese Methode bei gemischten Modellen i.a. keine erwartungstreuen Schätzer für die Varianzkomponenten liefert. Mit Hilfe einer geometrischen Darstellung der von S.R. SEARLE nur an Spezialfällen behandelten Methode III von HENDERSON ergibt sich ein Beispiel eines gemischten Modells, bei dem diese Schätzmethode versagt.

Im zweiten Kapitel wird zuerst C.R. RAO's "Minimum-Norm-Quadratic-Unbiased-Estimation" Methode zur Schätzung von Varianz- und Kovarianzkomponenten dargestellt und ihre Erwartungstreue im allgemeinen Fall bewiesen. Die von J. SEELY angegebenen Bedingungen für die Existenz von Schätzern, die in der Klasse aller erwartungstreuen

Schätzer gleichmäßig kleinste Varianz haben, erweisen sich als starke Forderungen an die Datenstruktur und zwingen zur Beschränkung auf diejenigen erwartungstreuen (translationsinvarianten) Schätzer, die invariant gegenüber einer Gruppe von Transformationen des Stichprobenraumes sind. In dieser Klasse von (translationsinvarianten) Schätzern wird dann ein Zusammenhang zwischen den Schätzmethoden von SEELY und RAO hergestellt.

Bei Vorliegen einer Spektraldarstellung der Kovarianzmatrix gibt P. BALTESTRA in [1] beste quadratische erwartungstreue Schätzer für die Eigenwerte der Kovarianzmatrix an. Es wird gezeigt, daß diese Methode bei einem Balanced Incomplete Block Design eine sehr zweckmäßige Angabe der von J. SEELY gefundenen optimalen Varianzkomponentenschätzer ermöglicht. Diese Angabe erlaubt die Anwendung der von P. HUBER in [4] ausgearbeiteten Theorie der robusten Schätzung in einem reinen Erwartungswertmodell, da die unbekannten Eigenwerte der Kovarianzmatrix mittels Kleinst-Quadrate-Schätzungen der Störvarianz eines modifizierten Regressionsmodells geschätzt werden.

Wie in einem einfach hierarchischen Versuchsplan gezeigt wird, stellt SEELY's Bedingung für die Existenz bester Schätzer in der Klasse der translationsinvarianten Schätzer eine sehr einschränkende Bedingung an die Parameter des Versuchsplanes. Deshalb werden die Spektraldarstellungen der Kovarianzmatrizen einer einfach- und zweifach-hierarchischen Klassifikation angegeben, um Schätzer nach BALESTRA's Methode zu berechnen.

In [2] hat H. DRYGAS bei einem reinen Erwartungswertmodell Schätzer für die Störvarianz σ^2 angegeben, die bezüglich der quadratischen Verlustfunktion ein kleineres GAUSS'sches Risiko haben

als der beste quadratische erwartungstreue Schätzer von σ^2 und in [3] setzt sich H. DRYGAS kritisch mit der erwartungstreuen Schätzung von Varianzkomponenten für ein Varianzkomponentenmodell mit Erwartungswert Null auseinander. Im letzten Abschnitt des zweiten Kapitels wird das GAUSS'sche Risiko einer quadratischen Form (zur Schätzung einer vorgegebenen Linearkombination von unbekannten Varianzkomponenten im gemischten Modell) berechnet und versucht, dieses gleichmäßig für alle unbekannten Modellparameter zu minimieren. Dabei zeigte sich, daß Schätzer, welche das GAUSS'sche Risiko gleichmäßig minimieren, nur in Sonderfällen angegeben werden können. Für eine einfache hierarchische Klassifikation und für das gemischte Modell eines Balanced Complete Block Designs werden verzerrte Schätzer berechnet, die ein kleineres GAUSS'sches Risiko liefern als die besten quadratischen erwartungstreuen Schätzer.

Im letzten Kapitel wird gezeigt, wie bei dem in der Ökonometrie (Behandlung von kombinierten Zeitreihen- und Querschnittsdaten) und Biometrie (Bestimmung von Wachstumskurven mittels wiederholter Messungen am selben Individuum) verwendeten Random-Regression-Coefficient-Modell die Schätzung der Kovarianzmatrix durchgeführt werden kann. Schließlich wird der mittels der geschätzten Kovarianzmatrix zu berechnende AITKEN-Schätzer für den unbekannten Parametervektor der festen Effekte angegeben.

Der Autor spricht Herrn Prof. Dr. V. BAUMANN für die vielen wissenschaftlichen Anregungen während der Betreuung und die Verbesserungsvorschläge bei der Fertigstellung dieser Arbeit seinen herzlichen Dank aus.

1. Orthogonalprojektoren und die Darstellung von klassischen Schätzmethoden für Varianzkomponenten

1.1. Modell und Problemstellung

Zur Beschreibung eines allgemeinen univariaten linearen Modells werden folgende Annahmen gemacht:

$n,m,q,s,l,k_{s+1},k_{s+2},\ldots,k_l$ sind natürliche Zahlen,
y ist ein n-dimensionaler Zufallsvektor mit der Darstellung

$$y = X\beta + \sum_{j=1}^{s} U_j v_j + \sum_{i=s+1}^{l} U_i v_i \qquad (1.1) \quad .$$

Dabei bedeuten X eine bekannte (n×m)-Matrix,

$\beta = (\beta_1,\beta_2,\ldots,\beta_m)' \in \mathbb{R}^m$ einen unbekannten Parametervektor ,

U_j eine bekannte (n×q)-Matrix für j = 1,2,...,s,

v_j einen q-dimensionalen Zufallsvektor mit $E(v_j) = 0$ und der unbekannten (q×q)-Kovarianzmatrix $\sum := E(v_j v_j')$ für j = 1,2,...,s,

U_i eine bekannte $(n\times k_i)$-Matrix für i = s+1,s+2,...,l,

v_i einen k_i-dimensionalen Zufallsvektor mit $E(v_i) = 0$ und $Cov(v_i) = \sigma_i^2 \cdot I_{k_i}$ mit unbekanntem $\sigma_i^2 > 0$ für i = s+1,s+2,...,l.

Sind die Zufallsvektoren v_k für k = 1,2,...,l außerdem paarweise unkorreliert, d.h. gilt $E(v_k v_{k'}') = 0$ für $1 \le k \ne k' \le l$, so ergibt sich aus (1.1) und den gemachten stochastischen Annahmen:

$$E(y) = X\beta \text{ und } Cov(y) = \sum_{j=1}^{s} U_j \sum U_j' + \sum_{i=s+1}^{l} \sigma_i^2 V_i$$

mit $V_i := U_i U_i'$ für $i = s+1, s+2, \ldots, l$.

Bei Vorliegen einer Realisation des Zufallsvektors y befaßt sich die Schätztheorie in linearen Modellen einerseits mit der Schätzung von schätzbaren Linearformen des unbekannten Parametervektors β und andererseits mit der Schätzung der unbekannten Kovarianzmatrix $\sum$ sowie der unbekannten Varianzkomponenten σ_i^2 für $i = s+1, s+2, \ldots, l$.

Die Darstellung (1.1) enthält folgende Spezialfälle:

(i) Ist $l = s$, so liegt ein reines Kovarianzkomponentenmodell vor. Gesucht ist eine Schätzung der $\frac{q}{2}(q+1)$ Elemente von $\sum$ (oder der Linearkombination $\mathrm{tr}(\sum \Lambda)$, wobei Λ eine gegebene symmetrische Matrix ist.)

(ii) Ist $s = 0$, so spricht man von einem gemischten Modell, in dem schätzbare Linearformen von β und die Varianzkomponenten σ_i^2 für $i = 1,2,\ldots,l$ zu schätzen sind.

(iii) Ist $s = 0$, $X = 1_n$ der n-dimensionale Spaltenvektor mit lauter Einsen und $\beta_1 = \beta_2 = \ldots = \beta_m =: \mu$ eine reelle Zahl, so handelt es sich um ein reines Varianzkomponentenmodell oder kurz: Modell II.

(iv) Ist $s = 0$, $l = 1$ und $\mathrm{Cov}(y) = \sigma^2 V$ mit bekannter symmetrischer positiv definiter $(n \times n)$-Matrix V und unbekanntem $\sigma^2 > 0$, so liegt ein reines Erwartungswertmodell oder kurz: Modell I vor.

Allen Modellen gemeinsam ist die geometrische Vorstellung, daß der Erwartungswertvektor E(y) des Zufallsvektors y in dem von den

Spaltenvektoren der Designmatrix X aufgespannten linearen Teilraum des $\mathbb{R}^n$ liegt.

Zur Darstellung der Schätztheorie der Kovarianzmatrix Cov(y) verwendet man Resultate aus der Theorie der endlich dimensionalen Vektorräume, welche im folgenden zusammengestellt werden.

1.2. Hilfsmittel, Orthogonalprojektoren und Erwartungswerte quadratischer Formen.

Ist A eine reelle (n×k)-Matrix, so nennt man den linearen Teilraum des $\mathbb{R}^n$, welcher von den Spaltenvektoren der Matrix A erzeugt wird, den Spaltenraum von A und bezeichnet ihn mit $\underline{R}(A)$. Ferner sei rg(A) der Rang von A, A' die zu A transponierte Matrix und $\underline{N}(A)$ der Nullraum von A. Ist n = k, so bedeutet tr A die Spur von A.

Im folgenden bezeichnet man für $a \in \mathbb{R}^n$ und $b \in \mathbb{R}^n$ mit (a,b) ein Skalarprodukt im $\mathbb{R}^n$ und nennt a orthogonal zu b, wenn (a,b) = 0 gilt. Die positive Quadratwurzel von (a,a) wird mit |a| abgekürzt und heißt die Länge von a. Ist T ein linearer Teilraum des $\mathbb{R}^n$, so ist $T^{\perp}$ die Menge aller Vektoren im $\mathbb{R}^n$, welche zu allen Vektoren in T orthogonal sind und heißt orthogonales Komplement von T.

Eine Menge von orthogonalen Vektoren der Länge 1 heißt eine Orthonormalbasis. Die Darstellung von Kovarianzmatrizen spezieller linearer Modelle wird erleichtert durch folgende

1.2.1. Definitionen:

(a) Sind A_1 bzw. A_2 zwei (p×q)-bzw. (r×s)-Matrizen, so heißt die unterteilte (p+r×q+s)-Matrix

$$A_1 \oplus A_2 := \left(\begin{array}{c|c} A_1 & \begin{matrix} 0 \\ (p\times s) \end{matrix} \\ \hline \begin{matrix} 0 \\ (r\times q) \end{matrix} & A_2 \end{array} \right) =: \bigoplus_{i=1}^{2} A_i$$ die direkte

Summe der Matrizen A_1 und A_2 . Entsprechend wird die Summe von mehr als zwei Matrizen gebildet.

(b) Unter dem direkten Produkt oder KRONECKER-Produkt der (p×q)-Matrix A und der (r×s)-Matrix B versteht man die unterteilte (pr×q·s)-Matrix

$$A \otimes B := \begin{pmatrix} a_{11}\cdot B & \dots & a_{1q}\cdot B \\ \vdots & & \vdots \\ a_{p1}\cdot B & \dots & a_{pq}\cdot B \end{pmatrix} .$$

Sei T ein linearer Teilraum des R^n und $T^\perp$ sein orthogonales Komplement, dann nennt man den Operator P, welcher auf T entlang $T^\perp$ projiziert, den Orthogonalprojektor auf T.

Die Eigenschaften eines Orthogonalprojektors sind angegeben in dem in [8;S.47] bewiesenen

1.2.2. Satz:

Sei durch (a,b) = a'Mb mit positiv definiter Matrix M ein Skalarprodukt im R^n gegeben. Dann ist P genau dann ein Orthogonalprojektor, wenn

(i) $P^2 = P$ und (ii) $MP = (MP)'$ gelten.

Für die Berechnung eines Orthogonalprojektors erweist sich als nützlich der

1.2.3. Satz über die Matrizendarstellung eines Orthogonalprojektors:

Gegeben seien $x_1,x_2,\ldots,x_k$ aus dem R^n mit $l := rg(x_1,x_2,\ldots,x_k)$ und $c_1,c_2,\ldots,c_l$ seien GRAM-SCHMIDT-Orthonormalisierte von $x_1,x_2,\ldots,x_k$ bezüglich des Skalarproduktes (a,b) = a'Mb. Dann ist $P := \sum_{i=1}^{l} c_i c_i'M$ der Orthogonalprojektor auf $\underline{R}(x_1,x_2,\ldots,x_k)$.

Beweis: $P^2 = \left(\sum_{i=1}^{l} c_i c_i'M\right)\left(\sum_{i=1}^{l} c_i c_i'M\right) = \sum_{i=1}^{l}\sum_{j=1}^{l} c_i c_i'M c_j c_j'M =$

$$\sum_{i=1}^{l} c_i c_i'M = P, \text{ da } c_i'Mc_j = \begin{cases} 1 & \text{falls } j = j \\ 0 & \text{falls } i \neq j \end{cases} \text{ gilt.}$$

$$MP = \sum_{i=1}^{l} Mc_i c_i'M = (MP)'.$$

Um zu zeigen, daß P der Orthogonalprojektor auf $\underline{R}(x_1,x_2,\ldots,x_k)$ ist, sei $y = y_1+y_2 \in R^n$ gegeben mit $y_1 \in \underline{R}(x_1,x_2,\ldots,x_k)$ und $y_2 \in \underline{R}(x_1,x_2,\ldots,x_k)^{\perp}$.

Dann gibt es reelle Zahlen λ_i für $i = 1,2,\ldots,l$ mit $y_1 = \sum_{i=1}^{l} \lambda_i c_i$ und $Py_1 = (\sum_{i=1}^{l} c_i c_i'M) \cdot \sum_{i=1}^{l} \lambda_i c_i = \sum_{i=1}^{l} \lambda_i c_i = y_1$, da $(c_i, \sum_{i=1}^{l} \lambda_i c_i) = c_i'M \cdot \sum_{i=1}^{l} \lambda_i c_i = \lambda_i$ gilt für $i = 1,2,\ldots,l$.

Ferner ist $Py_2 = (\sum_{i=1}^{l} c_i c_i'M)y_2 = 0$, da $y_2 \in \underline{R}(c_1,\ldots,c_l)^{\perp}$.

Die explizite Angabe eines Orthogonalprojektors macht Gebrauch von folgender

1.2.4. Definition: Sei A eine reelle (n×m)-Matrix. Eine verallgemeinerte Inverse von A oder kurz g-Inverse von A ist eine (m×n)-Matrix A^- mit der Eigenschaft, daß $x := A^-y$ eine Lösung der Gleichung $Ax = y$ für $y \in \underline{R}(A)$ ist.
In [8;S.24] wird gezeigt:

1.2.4'. Folgerung: A^- ist genau dann eine g-Inverse von A, wenn $AA^-A = A$ gilt.

Man verwendet im folgenden den in [8;S.47] angegebenen
1.2.5. Satz: Sei $\underline{R}(X)$ der von den Spaltenvektoren der (n×m)-Matrix X erzeugte Teilraum des R^n versehen mit dem Skalarprodukt $(a,b) = a'Mb$ mit positiv definiter Matrix M, dann ist der Orthogonalprojektor P auf $\underline{R}(X)$ gegeben durch
$P = X(X'MX)^-X'M$, wobei $(X'MX)^-$ eine g-Inverse von X'MX ist.

Bezeichnet tr(AB) die Spur des Matrizenproduktes der (r×c)-Matrix A mit der (c×r)-Matrix B, so gilt $tr(AB) = tr(BA)$.

Über die Berechnung des Erwartungswertes einer quadratischen Form gilt der in [9;S.55] bewiesene

1.2.6. Satz:

Sei y ein n-dimensionaler Zufallsvektor mit der Kovarianzmatrix Cov(y) und A eine symmetrische (n×n)-Matrix. Dann gilt

$E(y'Ay) = tr(A \cdot Cov(y)) + E(y)'AE(y)$.

1.3. HENDERSON's Methode I im reinen Varianzkomponentenmodell.

Gegeben seien n Beobachtungen des n-dimensionalen Zufallsvektors y mit der Darstellung $y = 1_n\mu + \sum_{i=1}^{l} U_i v_i$. (1.3.)

Bezeichnet man mit Q_m für m = 1,2,...,l-1 bzw. Q_o die Orthogonalprojektoren auf $\underline{R}(U_m)$ für m = 1,2,...,l-1 bzw. $\underline{R}(1_n)$ bezüglich des Skalarproduktes (a,b) = a'b, so läßt sich

HENDERSON's Methode I wie folgt darstellen:

1. Berechne $y'Q_jy$ für j = 0,1,...,l-1 bzw. y'y und
2. Löse das Gleichungssystem

$$(HI)\begin{cases} y'Q_jy = 1_n' \cdot \hat{\mu} Q_j 1_n \cdot \hat{\mu} + tr(Q_j \cdot \sum_{i=1}^{l} \hat{\sigma}_i^2 V_i) & (j = 0,1,\ldots,l-1) \\ y'y = 1_n' \cdot \hat{\mu} \cdot 1_n \cdot \hat{\mu} + tr(\sum_{i=1}^{l} \hat{\sigma}_i^2 V_i) \end{cases}$$

nach den Schätzern $\hat{\sigma}_i^2$ für die unbekannten Varianzkomponenten σ_i^2 für i = 1,2,...l auf.

Bezeichnet man lineare Modelle mit der Darstellung (1.3.), bei denen in jeder Zeile der Matrix U_i für $i = 1,2,\ldots,l$ genau einmal die Zahl 1 und sonst lauter Nullen stehen, als Varianzanalyse-Modell, so gilt folgender in [9;S. 424] an einem Beispiel erläuterter

1.3.1. Satz:

Handelt es sich bei einem reinen Varianzkomponentenmodell gleichzeitig um ein Varianzanalyse-Modell, so liefert HENDERSON's Methode I erwartungstreue Schätzer für die unbekannten Varianzkomponenten.

Beweis: Nach 1.2.6. ergibt sich

$$E(y'Q_jy) = \mu^2 \cdot 1_n'Q_j1_n + tr(Q_j \cdot \sum_{i=1}^{l} \sigma_i^2 V_i) \quad (j = 0,1,\ldots,l-1)$$

und $E(y'y) = \mu^2 1_n' \cdot 1_n + tr(\sum_{i=1}^{l} \sigma_i^2 V_i)$.

Bei einem Varianzanalyse-modell ist $1_n \in \underset{\sim}{R}(U_i)$ für $i = 1,2,\ldots,l$ und damit $1_n'\hat{\mu}Q_j1_n\hat{\mu} = n\cdot\hat{\mu}^2$ für $j = 0,1,\ldots,l-1$, da $Q_j1_n = 1_n$ für $j = 0,1,\ldots,l-1$.

Durch Subtraktion einer der $l+1$ Schätzgleichungen von den übrigen erhält man somit l Gleichungen, in denen nur noch die l unbekannten Schätzer für die zu schätzenden l Varianzkomponenten vorkommen.

Die Anwendung von HENDERSON's Methode I auf das

1.3.2. Beispiel

$$y = \begin{pmatrix} 100 \\ 200 \\ 300 \\ 400 \\ 800 \\ 600 \end{pmatrix} = \begin{pmatrix} 1 & 10 & 100 & 10 \\ 1 & 01 & 100 & 10 \\ 1 & 10 & 100 & 01 \\ 1 & 01 & 010 & 10 \\ 1 & 01 & 010 & 01 \\ 1 & 10 & 001 & 10 \end{pmatrix} \begin{pmatrix} \mu \\ v_{11} \\ v_{12} \\ v_{21} \\ v_{22} \\ v_{23} \\ v_{31} \\ v_{32} \end{pmatrix} + \begin{pmatrix} e_{111} \\ e_{211} \\ e_{112} \\ e_{221} \\ e_{222} \\ e_{131} \end{pmatrix} = 1_6 \cdot \mu + \sum_{i=1}^{4} U_i v_i$$

liefert folgendes System von Schätzgleichungen:

$$y'Q_o y = 96 \cdot 10^4 = 6\hat{\mu}^2 + 3\hat{\sigma}_1^2 + \frac{7}{3}\hat{\sigma}_2^2 + 4\hat{\sigma}_3^2 + \hat{\sigma}_4^2$$

$$y'Q_1 y = \frac{296}{3} \cdot 10^4 = 6\hat{\mu}^2 + 6\hat{\sigma}_1^2 + \frac{10}{3}\hat{\sigma}_2^2 + \frac{10}{3}\hat{\sigma}_3^2 + 2\hat{\sigma}_4^2$$

$$y'Q_2 y = 120 \cdot 10^4 = 6\hat{\mu}^2 + \frac{14}{3}\hat{\sigma}_2^2 + 6\hat{\sigma}_2^2 + \frac{11}{3}\hat{\sigma}_3^2 + 3\hat{\sigma}_4^2$$

$$y'Q_3 y = \frac{411}{4} \cdot 10^4 = 6\hat{\mu}^2 + 3\hat{\sigma}_1^2 + \frac{5}{2}\hat{\sigma}_2^2 + 6\hat{\sigma}_3^2 + 2\hat{\sigma}_4^2$$

$$y'y = 130 \cdot 10^4 = 6\hat{\mu}^2 + 6\hat{\sigma}_1^2 + 6\hat{\sigma}_2^2 + 6\hat{\sigma}_3^2 + 6\hat{\sigma}_4^2.$$

HENDERSON's Schätzer nach Methode I haben folgenden

1.3.3. Nachteil:

Methode I liefert bei gemischten Modellen Schätzer für die unbekannten Varianzkomponenten, welche i.a. von den festen Effekten abhängen.

Beweis: Nimmt man o.B.d.A. an, daß in dem reinen Varianzkomponentenmodell mit der Darstellung (1.3.)

v_1 ein Parametervektor von festen Effekten ist, welchen wir mit β bezeichnen, so ist mit $U_1 = X_1$

$$E(y) = 1_n \cdot \mu + X_1\beta.$$

Bezeichnet Q_o bzw. Q_m den Orthogonalprojektor auf $\underline{R}(1_n)$ bzw. $\underline{R}(U_m)$ für $m = 2,3,\ldots,l-1$, so ergibt sich aus dem Gleichungssystem (HI):

$$E(y'Q_jy) = [(1_n,X_1)(\mu,\beta')']'Q_j[(1_n,X_1)(\mu,\beta')'] + tr(Q_j \cdot \sum_{i=2}^{l} \sigma_i^2 V_i)$$

für $j = 0,2,3,\ldots,l-1$ und

$$E(y'y) = [(1_n,X_1)(\mu,\beta')']'[(1_n,X_1)(\mu,\beta')'] + tr(\sum_{i=2}^{l} \sigma_i^2 V_i).$$

Man erhält also l Gleichungen für die Schätzer von l-1 unbekannten Varianzkomponenten. Da

$$[(1_n,X_1)(\mu,\beta')']'Q_j[(1_n,X_1)(\mu,\beta')'] = n\cdot\mu^2 + 2\mu\cdot 1_n'X_1\beta + \beta'X_1'Q_jX_1\beta \text{ ist,}$$

elimiert man durch Subtraktion einer der l Gleichungen die Summe $n\cdot\mu^2 + 2\mu\cdot 1_n'X_1\beta$ aus obigem Gleichungssystem.

Da nun i.a. $X_1'Q_kX_1 \neq X_1'Q_hX_1$ ist für $k \neq h$, kommen in den restlichen l-1 Gleichungen für die Schätzer der l-1 Varianzkomponenten noch die festen Effekte β vor.

1.3.4. Beispiel: Bei folgendem gemischten Modell

$$y = \begin{pmatrix}1\\1\\1\\1\\1\end{pmatrix} \cdot \mu + \begin{pmatrix}10\\10\\10\\01\\01\end{pmatrix} \beta + \begin{pmatrix}100\\100\\001\\010\\001\end{pmatrix} v_2 + I_5 v_3 = 1_5 \cdot \mu + X_1\beta + U_2 v_2 + I_5 v_3$$

ergibt sich mit $\beta := (\beta_1, \beta_2)'$ das Gleichungssystem (HI):

$$y'Q_o y = 5\hat{\mu}^2 + 2\hat{\mu}(3\hat{\beta}_1 + 2\hat{\beta}_2) + \tfrac{1}{5}(9\hat{\beta}_1^2 + 12\hat{\beta}_1\hat{\beta}_2 + 4\hat{\beta}_2^2) + \tfrac{9}{5}\hat{\sigma}_2^2 + \hat{\sigma}_3^2$$

$$y'Q_2 y = 5\hat{\mu}^2 + 2\hat{\mu}(3\hat{\beta}_1 + 2\hat{\beta}_2) + \tfrac{1}{2}(5\hat{\beta}_1^2 + 2\hat{\beta}_1\hat{\beta}_2 + 4\hat{\beta}_2^2) + 5\hat{\sigma}_2^2 + 3\hat{\sigma}_3^2$$

$$y'y = 5\hat{\mu}^2 + 2\hat{\mu}(3\hat{\beta}_1 + 2\hat{\beta}_2) + 3\hat{\beta}_1^2 + 2\hat{\beta}_2^2 + 5\hat{\sigma}_2^2 + 5\hat{\sigma}_3^2.$$

Daraus erhält man

$$\tfrac{1}{2}\hat{\beta}_1^2 - \hat{\beta}_1\hat{\beta}_2 + 2\hat{\sigma}_3^2 = y'y - v'Q_2 y$$

und $-\tfrac{1}{5}\hat{\beta}_1^2 + \tfrac{2}{5}\hat{\beta}_1\hat{\beta}_2 - \tfrac{16}{5}\hat{\sigma}_2^2 = y'Q_o y - y'y$, d.h.

die Schätzer $\hat{\sigma}_2^2$ und $\hat{\sigma}_3^2$ sind von den festen Effekten abhängig.

1.4. HENDERSON's Methode III im gemischten Modell

Gegeben seien n Beobachtungen des Zufallsvektors y mit der Darstellung

$$y = X\beta + \sum_{i=1}^{l} U_i v_i \qquad (1.4.)$$

Mit P_i für i = 1,2,...,l bzw. P_o bezeichnet man die Orthogonalprojektoren auf $\underline{R}(X,U_1,U_2,\ldots,U_i)$ für i = 1,2,...,l bzw. $\underline{R}(X)$

bezüglich des Skalarproduktes $(a,b) = a'b$.

Ferner sei $N_k := P_{l-1} - P_{k-1}$ für $k = 1,2,\ldots,l-1$ und

$N_l := P_l - P_{l-1}$.

Mit diesen Bezeichnungen wird HENDERSON's Methode III wie folgt beschrieben:

1. Berechne $y'N_k y$ für $k = 1,2,\ldots,l$ und
2. löse das Gleichungssystem

(HIII) $y'N_k y = \sum_{i=k}^{l} \hat{\sigma}_i^2 tr(N_k V_i)$ $(k = 1,2,\ldots,l)$

nach den Schätzern $\hat{\sigma}_i^2$ für die unbekannten Varianzkomponenten σ_i^2 für $i = 1,2,\ldots,l$ auf.

Das in [9;S.446] für ein Beispiel angegebene Resultat lautet als

1.4.1. Satz

HENDERSON's Methode III liefert bei gemischten Modellen erwartungstreue Schätzer für die unbekannten Varianzkomponenten.

Beweis:

(i) Da $P_j X = X$ ist für alle $j = 0,1,\ldots,l$
gilt $N_i X = (P_{l-1} - P_{i-1})X = 0$ und $N_i X\beta = 0$
für alle β und alle $i = 1,2,\ldots,l$.

(ii) Da $Cov(y) = \sum_{i=1}^{l} \sigma_i^2 V_i$ ist, gilt

$$E(y'N_k y) = tr(N_k \cdot Cov(y)) + \beta'X'N_k X\beta = \sum_{i=1}^{l} \sigma_i^2 tr(N_k V_i)$$

für $k = 1,2,\ldots,l$.

Falls $i < k < l$ ist, gilt $N_k U_i = P_{l-1}U_i - P_{k-1}U_i = 0$, und damit $N_k V_i = N_k U_i U_i' = 0$.

Falls $i < k = l$ ist, gilt $N_k U_i = N_l U_i = P_l U_i - P_{l-1}U_i = 0$ und damit $N_k V_i = N_k U_i U_i' = 0$.

Damit ist

$$E(y'N_k y) = \sum_{i=k}^{l} \sigma_i^2 \operatorname{tr}(N_k V_i) \text{ gezeigt.}$$

1.4.2. Bemerkung über die Lösbarkeit des Gleichungssystems (HIII)

Schreibt man das Gleichungssystem (HIII) in der Form $\vec{R} = H\vec{\hat{\sigma}}^2$, wobei $\vec{R}$ den l-dimensionalen Vektor mit den Komponenten $y'N_k y$ für $k = 1,2,\ldots,l$ und $\vec{\hat{\sigma}}^2$ den l-dimensionalen Vektor der Schätzer der l unbekannten Varianzkomponenten bezeichnet, so ist (HIII) genau dann lösbar, und zwar eindeutig, wenn alle Diagonalelemente der Koeffizientenmatrix H von Null verschieden sind.

1.4.3. Bemerkungen über Nachteile von HENDERSON's Methode III:

(i) Die nach HENDERSON's Methode III ermittelten Schätzer hängen über die in (1.4.) festgesetzte Reihenfolge von der Numerierung der Varianzkomponenten ab.

(ii) Das folgende Beispiel zeigt, daß HENDERSON's Methode III nicht immer anwendbar ist:

Gegeben sei das Modell $y = X\beta + U_1 v_1 + U_2 v_2$ mit

$$X = \begin{pmatrix} 110 \\ 110 \\ 110 \\ 101 \\ 101 \end{pmatrix}, \quad U_1 = \begin{pmatrix} 1000 \\ 0100 \\ 0001 \\ 0010 \\ 0001 \end{pmatrix} \quad \text{und } U_2 = I_5.$$

Die Berechnung der in (HIII) vorkommenden Orthogonalprojektoren ergibt mit $J_n := 1_n 1_n'$
$P_o = \frac{1}{3}J_3 \oplus \frac{1}{2}J_2$, $P_1 = P_2 = I_5$ und damit $N_2 = 0$.

Das Gleichungssystem (HIII) ist nach 1.4.2. nicht eindeutig lösbar, da das Diagonalelement $tr(N_2V_2) = 0$ ist.

2. Neuere Methoden zur Schätzung von Kovarianzmatrizen und ihr Zusammenhang.

2.1. MINQUE-Theorie zur Schätzung von Kovarianzmatrizen

Man geht aus von der Darstellung des n-dimensionalen Beobachtungsvektors in der Form $y = X\beta+Uv$, wobei U bzw. v die Partition $U := (U_1,U_2,\ldots,U_l)$ bzw. $v := (v_1',v_2',\ldots,v_l')'$ bedeuten. Gesucht wird ein Schätzer $y'Ay$ für $\sum_{i=1}^{l} p_i\sigma_i$ mit gegebenem p_i für $i = 1,2,\ldots,l$ mit der Eigenschaft $AX = 0$. Die Bedingung $AX = 0$ ist notwendig und hinreichend dafür, daß für den Schätzer $y'Ay$ die Beziehung $y'Ay = (y+X\beta_o)'A(y+X\beta_o)$ für alle $\beta_o \in R^m$ gilt, d.h. der Wert des Schätzers soll sich nicht ändern, wenn sich die systematische Komponente von y um $X\beta_o$ mit beliebigem $\beta_o \in R^m$ ändert.

Die Forderung der Erwartungstreue führt zu

$$E(y'Ay) = \sum_{i=1}^{l} \sigma_i^2 \operatorname{tr}(AV_i), \text{ d.h. } \operatorname{tr}(AV_i) = p_i \text{ für } i = 1,2,\ldots,l.$$

RAO's Schätzmethode beruht auf folgender Überlegung:

Wäre der $\sum_{i=1}^{l} k_i$-dimensionale Zufallsvektor v ein beobachtbarer Zufallsvektor, so würde man $\sum_{i=1}^{l} p_i\sigma_i^2$ durch die Linearkombination der Stichprobenvarianzen $\sum_{i=1}^{l} p_i \cdot \frac{v_i'v_i}{k_i} =: v'\Delta v$ schätzen, wobei $\Delta := \bigoplus_{i=1}^{l} \frac{p_i}{k_i} \cdot I_{k_i}$ ist.

Die Differenz D zwischen dem aus den Beobachtungen y zu bestimmenden Schätzwert $y'Ay$ und $v'\Delta v$ ist $D := v'(U'AU-\Delta)v$.

2.1.1. Definition: Die quadratische Form y'Ay heißt MInimum Norm Quadratic Unbiased Estimator von $\sum_{i=1}^{l} \rho_i \sigma_i^2$, falls die Matrix A so bestimmt wird, daß die Norm $||U'AU-\Delta||$ der Matrix $U'AU-\Delta$ minimiert wird unter den Nebenbedingungen $AX = 0$ und $tr(AV_i) = \rho_i$ für $i = 1,2,\ldots,l$.

Bezeichnet $V := \sum_{k=1}^{l} V_k$ und wählt man als Matrixnorm die Summe der Quadrate der Matrizenelemente, so gilt folgender in [7;S.114] angegebene

2.1.2. Satz: Bezeichnet N_v den Orthogonalprojektor in $\underline{R}(X)^{\perp}$ bezüglich des Skalarproduktes $(a,b) = a'V^{-1}b$, so ist der MINQUE von $\sum_{i=1}^{l} \rho_i \sigma_i^2$ gegeben durch $A_M := \sum_{i=1}^{l} \rho_i V^{-1} N_v V_i V^{-1} N_v$, wobei die reellen Zahlen ρ_i so bestimmt werden, daß $\sum_{i=1}^{l} \rho_i tr(V^{-1} N_v V_i V^{-1} N_v V_j) = \rho_j$ für $j = 1,2,\ldots,l$ gilt.

Als Vorbereitung für die Schätzung von Varianz- und Kovarianzkomponenten im allgemeinen Fall sucht man jetzt eine Schätzung der $\frac{q}{2}(q+1)$ Elemente von Σ oder der Linearkombination $tr(\Sigma\Lambda)$ mit der vorgegebenen symmetrischen $(q \times q)$-Matrix Λ mittels einer quadratischen Form y'Ay mit der Eigenschaft $AX = 0$.

Mit $Cov(y) = \sum_{j=1}^{s} U_j \Sigma U_j'$ und $AX = 0$ gilt

$E(y'Ay) = tr(A \cdot Cov(y)) = tr(\sum \cdot \sum_{j=1}^{s} U_j'AU_j) = tr(\sum\Lambda)$.

Damit muß die Matrix A der Bedingung $\sum_{j=1}^{s} U_j'AU_j = \Lambda$ genügen.

Wären die q-dimensionalen Zufallsvektoren v_j beobachtbar, so würde man $tr(\sum\Lambda)$ durch $\frac{1}{s} \cdot \sum_{j=1}^{s} v_j'\Lambda v_j$ schätzen.

Der MINQUE von $tr(\sum\Lambda)$ ergibt sich als Lösung des Optimierungsproblems $y'Ay - \frac{1}{s} \cdot \sum_{j=1}^{s} v_j'\Lambda v_j = \min!$ unter den Nebenbedingungen $AX = 0$ und $\sum_{j=1}^{s} U_j'AU_j = \Lambda$. Damit gilt wie in [7;S.115]

2.1.3. Satz: Mit $V := \sum_{j=1}^{s} U_jU_j'$ wird dieses Optimierungsproblem gelöst durch $A_* := \sum_{i=1}^{s} V^{-1}N_vU_i\underline{P}U_i'V^{-1}N_v$, wobei $\underline{P}$ Lösung von $\sum_{j=1}^{s}\sum_{i=1}^{s} U_j'V^{-1}N_vU_i\underline{P}U_i'V^{-1}N_vU_j = \Lambda$ ist.

Nun berechnet man die Schätzung von Varianz- und Kovarianzkomponenten im allgemeinen Fall wie folgt:

Unter Verwendung der Abkürzungen $V := \sum_{i=1}^{l} U_iU_i'$ und $N_v := I-X(X'V^{-1}X)^-X'V^{-1}$ lautet der in [7;S.115] ohne Beweis angegebene

2.1.4. Satz: Das folgende Gleichungssystem (1) und (2) liefert im allgemeinen Fall erwartungstreue Schätzer $\hat{\sum}, \hat{\sigma}^2_{s+1}, \ldots, \hat{\sigma}^2_l$ für $\sum, \sigma^2_{s+1}, \ldots, \sigma^2_l$.

$$(1)\quad \sum_{i=1}^{s} U_i'V^{-1}N_v yy'V^{-1}N_v U_i = \sum_{i=1}^{s}\sum_{j=1}^{s} U_i'V^{-1}N_v U_j \hat{\Sigma} U_j'V^{-1}N_v U_i +$$

$$+ \sum_{j=s+1}^{l} \hat{\sigma}_j^2 \sum_{i=1}^{s} U_i'V^{-1}N_v U_j U_j'V^{-1}N_v U_i$$

und

$$(2)\quad y'V^{-1}N_v U_j U_j'V^{-1}N_v y = \sum_{i=1}^{s} \operatorname{tr}(U_j'V^{-1}N_v U_i \hat{\Sigma} U_i'V^{-1}N_v U_j) +$$

$$+ \sum_{i=s+1}^{l} \hat{\sigma}_i^2 \operatorname{tr}(U_j'V^{-1}N_v U_i U_i'V^{-1}N_v U_j)$$

für $j = s+1, s+2, \ldots, l$.

Beweis: Mit $E(yy') = \operatorname{Cov}(y) + E(y)\cdot E(y)'$ und $N_v E(y) = 0$ gilt

$$E\Big[\sum_{i=1}^{s} U_i'V^{-1} N_v yy'V^{-1}N_v U_i\Big] = \sum_{i=1}^{s} U_i'V^{-1}N_v \Big[\sum_{j=1}^{s} U_j \Sigma U_j' + \sum_{j=s+1}^{l} \sigma_j^2 U_j U_j'\Big] V^{-1}N_v U_i$$

$$= \sum_{i=1}^{s}\sum_{j=1}^{s} U_i'V^{-1}N_v U_j \Sigma U_j'V^{-1}N_v U_i + \sum_{j=s+1}^{l} \sigma_j^2 \cdot \sum_{i=1}^{s} U_i'V^{-1}N_v U_j U_j'V^{-1}N_v U_i .$$

$$E(y'V^{-1}N_v U_j U_j'V^{-1}N_v y) =$$

$$= \operatorname{tr}\Big[V^{-1}N_v U_j U_j'V^{-1}N_v \cdot \Big(\sum_{i=1}^{s} U_i \Sigma U_i' + \sum_{i=s+1}^{l} \sigma_i^2 U_i U_i'\Big)\Big] =$$

$$= \sum_{i=1}^{s} \operatorname{tr}(U_j'V^{-1}N_v U_i \Sigma U_i'V^{-1}N_v U_j) + \sum_{i=s+1}^{l} \sigma_i^2 \cdot \operatorname{tr}(U_j'V^{-1}N_v U_i U_i'V^{-1}N_v U_j).$$

Bemerkung: Geschätzt werden $\frac{q}{2}(q+1)$ Elemente der unbekannten Kovarianzmatrix Σ und l-s unbekannte Varianzkomponenten $\sigma^2_{s+1},\sigma^2_{s+2},\ldots,\sigma^2_l$.
Dazu liefert Gleichung (1) $\frac{q}{2}(q+1)$ Gleichungen und Gleichung (2) l-s Gleichungen.

2.2. SEELY's Methode und ihr Zusammenhang mit der MINQUE-Theorie

Im folgenden heißt die Linearkombination $\sum_{i=1}^{l} p_i\sigma_i^2$ mit $p_i \in \mathbb{R}$ für $i = 1,2,\ldots,l$ in dem gemischten linearen Modell genau dann quadratisch schätzbar, wenn es eine reelle, symmetrische Matrix A gibt mit $E(y'Ay) = \sum_{i=1}^{l} p_i\sigma_i^2$. Ferner nennt man y'By genau dann einen besten erwartungstreuen quadratischen Schätzer für $\sum_{i=1}^{l} p_i\sigma_i^2$, wenn

(1) $E(y'By) = \sum_{i=1}^{l} p_i\sigma_i^2$ gilt und

(2) aus $E(y'Ay) = \sum_{i=1}^{l} p_i\sigma_i^2$ stets $var(y'By) \leq var(y'Ay)$ folgt.

Bezeichnet man mit N den Orthogonalprojektor in $\underline{R}(X)^{\perp}$ bezüglich des Skalarproduktes $(a,b) = a'b$, so gilt nach [10;S.1742] folgender

2.2.1. Satz:

Bei einem gemischten linearen Modell ist $\sum_{i=1}^{l} p_i\sigma_i^2$ genau dann

quadratisch schätzbar, wenn das lineare Gleichungssystem

(S) $\sum_{i=1}^{l} \operatorname{tr}(V_j N V_i)\rho_i = p_j$ $(j = 1,2,\ldots l)$ lösbar ist.

Der erwartungstreue Schätzer für $\sum_{i=1}^{l} p_i\sigma_i^2$ ist nach [10;S.1743] angegeben in

2.2.2. Korollar:

Erfüllen ρ_i und p_j für $i,j = 1,2,\ldots,l$ das Gleichungssystem (S) von 2.2.1., so ist $\sum_{i=1}^{l} \rho_i y'NV_i y$ ein erwartungstreuer Schätzer von $\sum_{i=1}^{l} p_i\sigma_i^2$.

Durch $y_{ijk} = \mu + \beta_i + c_j + s_{ij} + e_{ijk}$ für $i = 1,2,\ldots,a$; $j = 1,2,\ldots,b$ und $k = 0,1,2,\ldots,n_{ij}$ ist ein gemischtes lineares Modell mit den festen Effekten μ und β_i für $i = 1,2,\ldots,a$ gegeben ("Kreuzklassifikation" mit "Interaktionen" s_{ij}).

Mit $n := \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}$, $n_{i\cdot} := \sum_{j=1}^{b} n_{ij}$, $n_{\cdot j} := \sum_{i=1}^{a} n_{ij}$,

$N := I - \bigoplus_{i=1}^{a} \frac{1}{n_{i\cdot}} J_{(n_{i\cdot}\times n_{i\cdot})}$, $z_{ij} := \operatorname{tr}(V_j N V_i)$ als (i,j)-tem Element der Matrix Z, $\rho' := (\rho_1\ \rho_2,\ldots,\rho_l)$ und $p' := (p_1,p_2,\ldots,p_l)$

lautet das Gleichungssystem (S): $Z\rho = p$ mit

$$Z = \begin{pmatrix} \sum_{j=1}^{b} n_{\cdot j}^2 - \sum_{j=1}^{b} n_{\cdot j}\left(\sum_{i=1}^{a} \frac{n_{ij}^2}{n_{i\cdot}}\right) & \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}^2 - \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{n_{ij}^3}{n_{i\cdot}} & n - \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{n_{ij}^2}{n_{i\cdot}} \\ \sum_i\sum_j n_{ij}^2 - \sum_i\sum_j \frac{n_{ij}^3}{n_{i\cdot}} & \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}^2 - \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{n_{ij}^3}{n_{i\cdot}} & n - \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{n_{ij}^2}{n_{i\cdot}} \\ n - \sum_i\sum_j \frac{n_{ij}^2}{n_{i\cdot}} & n - \sum_i\sum_j \frac{n_{ij}^2}{n_{i\cdot}} & n - a \end{pmatrix}$$

Um eine hinreichende Bedingung dafür anzugeben, daß der in 2.2.2. angegebene Schätzer für $\sum_{i=1}^{l} p_i\sigma_i^2$ gleichmäßig kleinste Varianz unter allen erwartungstreuen Schätzern für $\sum_{i=1}^{l} p_i\sigma_i^2$ hat, verwendet man folgende

2.2.3. Definition:

Ein Teilraum $\mathfrak{L}$ des Raumes $\mathfrak{A}$ aller reellen, symmetrischen (n×n)-Matrizen heißt ein quadratischer Teilraum von $\mathfrak{A}$, wenn aus $B \in \mathfrak{L}$ folgt $B^2 \in \mathfrak{L}$.

Eine notwendige und hinreichende Bedingung für das Vorliegen eines quadratischen Teilraumes liefert das in [11;S.711] angegebene

2.2.4. Lemma:

Sei $\mathfrak{L}$ ein Teilraum von $\mathfrak{A}$ und $\mathfrak{L}_1$ eine $\mathfrak{L}$ erzeugende Menge von Matrizen aus $\mathfrak{A}$, dann ist $\mathfrak{L}$ genau dann ein quadratischer Teilraum von $\mathfrak{A}$, wenn aus $A,B \in \mathfrak{L}_1$ folgt $(A+B)^2 \in \mathfrak{L}$.

Der in [11;S.716] angegebene Satz lautet:

2.2.5. Satz:

Sei y ein n-dimensionaler Zufallsvektor, welcher normalverteilt ist mit $E(y) = X\beta$ und $Cov(y) = \sum_{i=1}^{l} \sigma_i^2 V_i$, gelten

(a) $sp\{V_1, V_2, \ldots V_l\}$ ist ein quadratischer Teilraum von $\mathfrak{A}$ und

(b) $\underline{R}(X)$ ist invarianter Teilraum von V_i, d.h. $V_i\underline{R}(X) \subset \underline{R}(X)$ für $i = 1,2,\ldots,l$,

dann hat der Schätzer $\sum_{i=1}^{l} \rho_i y'NV_iNy$ gleichmäßig kleinste Varianz unter allen erwartungstreuen Schätzern für $\sum_{i=1}^{l} p_i\sigma_i^2$, falls die ρ_i's dem Gleichungssystem (S) von 2.2.1. genügen.

Wie stark die Voraussetzung (b) in 2.2.5. ist, wird deutlich an folgendem

2.2.6. Beispiel:

Durch $y_{ijk} = \mu + \beta_i + c_j + e_{ijk}$ für $i = 1,2,\ldots,a$; $j = 1,2,\ldots,b$ und $k = 0,1,2,\ldots,n_{ij}$ ist ein gemischtes lineares Modell mit den festen Effekten μ und β_i für $i = 1,2,\ldots,a$ gegeben. Die in der Matrizendarstellung von Cov(y) vorkommenden Matrizen V_1 und V_2 sind gegeben durch

$$V_1 = \begin{pmatrix} \bigoplus_{j=1}^{b} J_{(n_{1j}\times n_{1j})} & \bigoplus_{j=1}^{b} J_{(n_{1j}\times n_{2j})} & \cdots & \bigoplus_{j=1}^{b} J_{(n_{1j}\times n_{aj})} \\ \cdot & \bigoplus_{j=1}^{b} J_{(n_{2j}\times n_{2j})} & \cdots & \bigoplus_{j=1}^{b} J_{(n_{2j}\times n_{aj})} \\ \cdots & \cdots & \ddots & \vdots \\ \cdot & \cdot & & \bigoplus_{j=1}^{b} J_{(n_{aj}\times n_{aj})} \end{pmatrix} \text{ und } V_2 = I_n.$$

Da $X = (1_n, \overset{a}{\underset{i=1}{\Sigma}} \oplus 1_{n_{i.}})$ mit $n_{i.} := \overset{b}{\underset{j=1}{\Sigma}} n_{ij}$ gilt, ist das Bild B_{i+1} des (i+1)-ten Spaltenvektors der Matrix X für i = 1,2,...,a unter V_1 gegeben durch $B_{i+1} =$

$$\underbrace{(n_{i1},\ldots,n_{i1}}_{n_{11}}, \underbrace{n_{i2},\ldots,n_{i2}}_{n_{12}},\ldots, \underbrace{n_{i1},\ldots,n_{i1}}_{n_{a1}},\ldots, \underbrace{n_{ia},\ldots,n_{ia}}_{n_{ab}})'$$

und damit $V_1\underline{R}(X) \subset \underline{R}(X)$ genau dann, wenn n_{ij} gleich ist für j = 1,2,...,b. Das heißt, die Voraussetzung (b) in 2.2.5. verlangt, daß der i-te feste Effekt β_i mit jedem zufälligen Effekt c_j gleich oft kombiniert wird. Da der in 2.2.5. angegebene Schätzer nur bei sehr spezieller Datenstruktur in der Gesamtheit aller erwartungstreuen Schätzer gleichmäßig kleinste Varianz hat, beschränkt man sich auf diejenige Teilgesamtheit der erwartungstreuen quadratischen Schätzer, die invariant gegenüber der Gruppe

$$\mathcal{G} := \{g: g(y) = y + r,\ r \in \underline{R}(X)\}$$

von Transformationen g des Stichprobenraumes ist.

Ist Q eine (n×q)-Matrix mit q = n - rgX, $\underline{R}(Q) = \underline{N}(X')$ und $Q'Q = I_q$, dann ist nach [12;S.1646] Q'y eine gegenüber $\mathcal{G}$ maximalinvariante Statistik von R^n auf R^q.

2.2.7. Hilfssatz:

Eine quadratische Form y'Ay ist genau dann invariant gegenüber $\mathcal{G}$, wenn AX = 0 gilt.

Beweis: (i) Gilt AX = 0, dann ist $(y+X\beta_o)'A(y+X\beta_o) = y'Ay$ für alle $y \in R^n$ und alle $\beta_o \in R^m$.

(ii) Ist y'Ay invariant gegenüber $\mathcal{C}$, dann folgt aus
$(y+X\beta_o)'A(y+X\beta_o) = y'Ay$ für alle $y \in R^n$ und alle $\beta_o \in R^m$, daß $2y'AX\beta_o + \beta_o'X'AX\beta_o = 0$ gilt für alle $y \in R^n$ und alle $\beta_o \in R^m$, d.h. $y'AX\beta_o = 0$ für alle $\beta_o \in R^m$ und alle $y \in R^n$ und daher $AX = 0$.

Ist y ein n-dimensionaler Beobachtungsvektor eines gemischten linearen Modells, dann ist $E(Q'y) = 0$ und $Cov(Q'y) = \sum_{i=1}^{l} \sigma_i^2 W_i$, wobei $W_i := Q'V_iQ$ ist für $i = 1,2,\ldots,l$.

Für quadratische Schätzer y'Ay mit $AX = 0$ für $\sum_{i=1}^{l} p_i\sigma_i^2$ ergibt sich aus 2.2.1. die

2.2.8. Folgerung:

$\sum_{i=1}^{l} p_i\sigma_i^2$ ist genau dann quadratisch schätzbar mit $AX = 0$, wenn das Gleichungssystem $\sum_{k=1}^{l} tr(W_kW_i)\rho_k = p_i$ für $i = 1,2,\ldots,l$ lösbar ist. Dann ist $\sum_{k=1}^{l} \rho_k y'QW_kQ'y$ ein erwartungstreuer Schätzer für $\sum_{i=1}^{l} p_i\sigma_i^2$.

Beweis: Wendet man 2.2.1. auf das gemischte lineare Modell $Q'y = \sum_{i=1}^{T} Q'U_iv_i$ mit $Cov(Q'y) = \sum_{i=1}^{l} \sigma_i^2 W_i$ an, so ist der Orthogonalprojektor in $\underset{q\times m}{R}\begin{pmatrix} 0 \\ \end{pmatrix}^{\perp}$ gegeben durch I_q.

Damit ist die Lösbarkeit des Gleichungssystems

$$\sum_{i=1}^{l} tr(Q'V_jQ\, I_q\, Q'V_iQ)\rho_i = \sum_{i=1}^{l} tr(W_jW_i)\rho_i = p_j \text{ für } j = 1,2,\ldots,l$$

notwendig und hinreichend für die quadratische Schätzbarkeit von $\sum_{i=1}^{l} p_i\sigma_i^2$ mit $AX = 0$.

Ferner ist nach 2.2.2. $\sum_{k=1}^{l} \rho_k y'QW_kQ'y$ ein erwartungstreuer Schätzer für $\sum_{i=1}^{l} p_i\sigma_i^2$.

Eine hinreichende Bedingung dafür, daß es in der Gesamtheit aller erwartungstreuen Schätzer, welche Funktionen des Zufallsvektors Q'y sind, Schätzer mit gleichmäßig kleinster Varianz gibt, wird nach [11;S.718] genannt durch

2.2.9. Satz:

Sei y ein n-dimensionaler Zufallsvektor, welcher normalverteilt ist mit $E(y) = X\beta$ und $Cov(y) = \sum_{i=1}^{l} \sigma_i^2 V_i$. Ferner sei die lineare Hülle von $\{Q'V_1Q, Q'V_2Q,\ldots,Q'V_lQ\}$ ein quadratischer Teilraum der Menge $\mathcal{C}$ aller reellen, symmetrischen $(q\times q)$-Matrizen und $\sum_{i=1}^{l} p_i\sigma_i^2$ sei quadratisch schätzbar mit $AX = 0$. Damit ist $\sum_{k=1}^{l} \alpha_k y'NV_kNy$ der beste Schätzer in der Gesamtheit aller erwartungstreuen Schätzer, die Funktionen von Q'y sind. Dabei gilt

$$\sum_{j=1}^{l} tr(V_iNV_jN)\alpha_j = p_i \text{ für } i = 1,2,\ldots,l.$$

Zur Herstellung des Zusammenhanges mit der MINQUE-Theorie verwendet man den in [8;S.65] angegebenen

2.2.10. Hilfssatz

Ist A eine symmetrische (n×n)-Matrix, $P := X(X'X)^{-}X'$ der Orthogonalprojektor in $\underline{R}(X)$ bezüglich des Skalarproduktes (a,b) = a'b und N = I - P, so ist eine Lösung $\tilde{A}$ des Optimierungsproblems $tr(A)^2 = \min!$ unter den Nebenbedingungen AX = 0 und $tr(AV_i) = p_i$ für i = 1,2,...,l gegeben durch $\tilde{A} = \sum_{k=1}^{l} \alpha_k NV_kN$, wobei die α's Lösungen des Gleichungssystems $\sum_{k=1}^{l} tr(NV_kNV_j)\alpha_k = p_j$ für j = 1,2,...,l sind.

2.2.11. Folgerung:

Ersetzt man in 2.1.2. das Skalarprodukt (a,b) = $a'V^{-1}b$ durch a'b, so stimmt der dort angegebene MINQUE für $\sum_{i=1}^{l} p_i\sigma_i^2$ mit den in 2.2.8. angegebenen Schätzern überein.

Beweis: Da QQ' = N ist, ist die Lösbarkeit des Gleichungssystems $\sum_{i=1}^{l} tr(NV_iNV_j)\rho_i = p_j$ für j = 1,2,...,l wegen

$$tr(QQ'V_iQQ'V_j) = tr(Q'V_iQQ'V_jQ) = tr(W_iW_j)$$

gleichwertig mit der Lösbarkeit des in 2.2.8. angegebenen Gleichungssystems.

Ferner ist $A = \sum_{k=1}^{l} \alpha_k NV_kN = \sum_{k=1}^{l} \alpha_k QQ'V_kQQ' =$

$$= \sum_{k=1}^{l} \alpha_k QW_kQ'.$$

Bevor SEELY's Resultate auf spezielle Designs angewandt werden, wird BALESTRA's Schätztheorie dargestellt.

2.3. Weitere Zusammenhänge neuerer Schätzmethoden

Gegeben sei ein lineares Modell mit der Darstellung $y = X\beta + e$, wobei X eine gegebene $(n \times m)$-Matrix, β ein unbekannter Parametervektor und e ein n-dimensionaler Zufallsvektor ist, der normalverteilt ist mit $E(e) = 0$ und $\mathrm{Cov}(e) = \sum_{h=1}^{p} \lambda_h C_h C_h' = \sum_{h=1}^{p} \lambda_h M_h$.

Dabei bezeichne λ_h für $h = 1,2,\ldots,p$ die verschiedenen Eigenwerte von Cov(e), C_h die zugehörigen Matrizen, deren Spalten aus den zum Eigenwert λ_h gehörigen Eigenvektoren bestehen, und $M_h := C_h C_h'$ die idempotenten Matrizen mit $M_i M_j = 0$ für $i \neq j$ und $\sum_{h=1}^{p} M_h = I_n$.

Für lineare Modelle dieser Art gilt folgender in [1;S.27] angegebener

2.3.1. Satz:

Der beste quadratische erwartungstreue Schätzer von λ_h ist gegeben durch die mittels der Methode der kleinsten Quadrate geschätzten Störvarianz der durch Hinzunahme von dummy-Variablen modifizierten Darstellung $y = Z_h \alpha_h + e$ mit $Z_h := (C_1, C_2, \ldots, C_{h-1}, C_{h+1}, C_{h+2}, \ldots, C_p, X)$.

Bemerkung:

λ_h ist nur schätzbar, wenn die Vielfachheit von λ_h größer als die Anzahl m der Spalten der Matrix X ist, denn sonst hat der

Modellraum $\underline{R}(Z_h)$ der modifizierten Darstellung eine höhere Dimension als der Beobachtungsraum R^n.

2.3.1'. Folgerung:

Unter den Voraussetzungen von 2.3.1. ist der Schätzer von λ_h gegeben

durch $\hat{\lambda}_h = \dfrac{y'(M_h - M_h X(X'M_h X)^{-1} X'M_h)\, y}{rgM_h - rgM_h X}$.

Beweis:

(i) Nach [8;S.33] gilt für symmetrische Matrizen A und D

$$\begin{vmatrix} A & B \\ B' & D \end{vmatrix}^{-1} = \begin{vmatrix} A^{-1} + FE^{-1}F' & -FE^{-1} \\ -E^{-1}F' & E^{-1} \end{vmatrix} \quad \text{mit}$$

$E = D - B'A^{-1}B$ und $F = A^{-1}B$.

(ii) Wegen 2.3.1. ist $\hat{\lambda}_h = \dfrac{y'(I_n - Z_h(Z_h'Z_h)^{-1}Z_h')y}{n - rgZ_h}$,

und man muß

$M_h - M_h X(X'M\,X)^{-1}X'M_h = I_n - Z_h(Z_h'Z_h)^{-1}Z_h'$ zeigen.

Dies erfolgt mit Hilfe von (i) unter Verwendung von $E^{-1} = (X'M_h X)^{-1}$ und den Eigenschaften der Spektraldarstellung von Cov(e).

(III) Da die Matrizen $Z_h(Z_h'Z_h)^{-1}Z_h'$, M_h und $M_h X(X'M_h X)^{-1}X'M_h$ idempotent sind, ergibt sich

$$rg(I_n - Z_h(Z_h'Z_h)^{-1}Z'_h) = n - rgZ_h =$$

$$rg(M_h - M_hX(X'M_hX)^{-1}X'M_h) = rg\ M_h - rg\ M_hX.$$

Eine formale Beschreibung spezieller Versuchspläne liefern die in [5;S.54,S.368] gegebenen

2.3.2. Definitionen:

(a) Ein Balanced Incomplete Block Design, (kurz: BIB-Design) ist eine Anordnung von t Symbolen("Behandlungen") in b Mengen ("Blöcken"), wobei jede Menge k Symbole enthält und folgende Bedingungen gelten:

(i) Jedes Symbol kommt höchstens einmal in einer Menge vor,

(ii) Jedes Symbol kommt in genau r Mengen vor,

(iii) Jedes Paar von Symbolen kommt zusammen in λ Mengen vor.

(b) Ein BIB-Design heißt symmetrisch, wenn t = b ist.

(c) Ein BIB-Design heißt ein Balanced Complete Block Design, wenn k = t ist.

Im folgenden sei die Beobachtung y_{ij} beschrieben durch

$$y_{ij} = \mu + \beta_i + b_j + e_{ij} \text{ für } i = 1,2,\ldots,t \text{ und } j = 1,2,\ldots,b$$

oder in Matrixform durch

$$y = X\beta + U_1v_1 + U_2v_2,$$

wobei $\beta := (\mu,\beta_1,\ldots,\beta_t)'$ ein unbekannter Parametervektor, $v_1 := (b_1,b_2,\ldots,b_b)'$ und $v_2 := (e_{11},e_{12},\ldots,e_{tb})'$ normalverteilte Zufallsvektoren sind mit $E(v_1) = 0$, $E(v_2) = 0$, $Cov(v_1) = \sigma_1^2 I_b$, $Cov(v_2) = \sigma_2^2 I_n$ und $E(v_1v_2') = 0$.

Genügt die Datenstruktur den in 2.3.2. angegebenen Bedingungen eines Balanced Complete Block Designs, dann ist $X = (1_n, I_t \otimes 1_b)$,

$$U_1 = 1_t \otimes I_b \text{ und } U_2 = I_{t \cdot b}$$

Da $U_1'U_1X = (t \cdot 1_n, J_{(n \times t)}) = V_1X$ gilt, ist $V_1\underline{R}(X) \subset \underline{R}(X)$ und damit ist Voraussetzung (b) von 2.2.5. erfüllt. Ferner ist $V_1^2 = t \cdot V_1$ und damit $sp\{V_1, V_2\}$ ein quadratischer Teilraum von $\mathcal{A}$.

Der in 2.2.5. angegebene Schätzer hat somit gleichmäßig kleinste Varianz unter allen erwartungstreuen Schätzern für $\sum_{i=1}^{l} p_i\sigma_i^2$.

Um im Fall $k < t$ 2.2.9. anzuwenden, muß man zeigen, daß die lineare Hülle von $\{Q'V_1Q, I_q\}$ ein quadratischer Teilraum von $\mathcal{C}$ ist. Eine notwendige und hinreichende Bedingung dafür liefert folgender in [11;S.719] ohne Beweis angegebener

2.3.3. Satz:

Die lineare Hülle von $\{Q'V_1Q, I_q\}$ ist genau dann ein quadratischer Teilraum von $\mathcal{C}$, wenn die Matrix $Q'V_1Q$ nicht mehr als zwei verschiedene Eigenwerte hat.

Beweis:

1. Gibt es reelle Zahlen ξ und η mit $(Q'V_1Q)^2 = \xi I_q + \eta Q'V_1Q$ und seien $d_1, d_2, \ldots, d_k$ die verschiedenen Eigenwerte von $Q'V_1Q$, so gilt $\xi + \eta \cdot d_i = d_i^2$ für $1,2,\ldots,k$. Da aber ξ und η eindeutig sind, ist $k \leq 2$.

2. Hat $Q'V_1Q$ nicht mehr als zwei verschiedene Eigenwerte, so

gibt es eine orthogonale Matrix P und zwei ganze Zahlen q_1 und q_2 mit $q_1 + q_2 = q$, so daß $Q'V_1Q = P(d_1 I_{q_1} \oplus d_2 I_{q_2})P'$ gilt.

Damit ist $(Q'V_1Q)^2 = (d_1+d_2)Q'V_1Q + d_1 \cdot d_2 I_q$.

Aus 2.2.9. ergibt sich die in [11;S.720] genannte

2.3.4. Folgerung:

Im Fall b = t hat $Q'V_1Q$ als einzigen von Null verschiedenen Eigenwert $\lambda_1^* := \frac{\lambda \cdot t}{r}$ mit der Vielfachheit t-1 und mit f := n-2t+1 hat $\frac{\lambda t\delta_2 - r\delta_1}{\lambda \cdot t \cdot f}\, y'(N - \frac{r}{\lambda \cdot t} NV_1N)y + \frac{\delta_1}{q} y'(\frac{r}{\lambda \cdot t} NV_1N)y$

unter allen erwartungstreuen Schätzern für $\delta_1\sigma_1^2 + \delta_2\sigma_2^2$, welche Funktionen von Q'y sind, kleinste Varianz.

2.3.5. Satz:

Die Spektraldarstellung von Cov(Q'y) ist im Falle b = t gegeben durch $Cov(Q'y) = (\sigma_1^2\lambda_1^* + \sigma_2^2)M_1^* + \sigma_2^2 M_2^*$,

wobei $M_1^* = \frac{Q'V_1Q}{\lambda_1^*}$ und $\lambda_1^* = \frac{\lambda \cdot t}{r}$ ist.

Beweis: Die Spektraldarstellung von $Q'V_1Q$ ist nach 2.3.4.

$$Q'V_1Q = \lambda_1^* M_1^* + 0 \cdot M_2^* \text{ mit } M_2^* = I_q - M_1^*.$$

Damit ist $Cov(Q'y) = \sigma_1^2\, Q'V_1Q + \sigma_2^2 Q'V_2Q = \sigma_1^2 Q'V_1Q + \sigma_2^2 I_q$

$$= \sigma_1^2\, \lambda_1^* M_1^* + \sigma_2^2 I_q$$

$$= (\sigma_1^2\lambda_1^* + \sigma_2^2)M_1^* + \sigma_2^2 M_2^*$$

Ferner ist

$$M_1^{*2} = \frac{Q'V_1Q}{\lambda_1^*} \cdot \frac{Q'V\,Q}{\lambda_1^*} = P(I_{t-1} \oplus \begin{smallmatrix} 0 \\ q-t+1 \end{smallmatrix})P'P(I_{t-1} \oplus \begin{smallmatrix} 0 \\ q-t+1 \end{smallmatrix})P'$$

$$= \frac{Q'V_1Q}{\lambda_1^*} = M_1^* \text{ mit orthogonaler Matrix P,}$$

$$M_2^{*2} = (I_q - M_1^*)(I_q - M_1^*) = I_q - M_1^* - M_1^* + M_1^{*2} = I_q - M_1^* = M_2^*$$

und $M_1^* \cdot M_2^* = \frac{Q'V_1Q}{\lambda_1^*}(I_q - \frac{Q'V_1Q}{\lambda_1^*}) = 0.$

2.3.6. Folgerung:

Sind M_1^* und M_2^* die idempotenten und symmetrischen Matrizen der Spektraldarstellung von Cov(Q'y), so gilt mit der Abkürzung $\frac{1}{f}(\delta_2 - \frac{1}{\lambda_1^*}\delta_1) =: a_1$ und $\frac{\delta_1}{q} =: a_2$ die Beziehung

$$\frac{1}{f}(\delta_2 - \frac{1}{\lambda_1^*}\delta_1)y'(N - \frac{1}{\lambda_1^*}NV_1N)y + \frac{\delta_1}{q}y'(\frac{1}{\lambda_1^*}NV_1N)y =$$

$$= a_1y'QM_2^*Q'y + a_2y'QM_1^*Q'y.$$

Beweis: 1. $N - \frac{1}{\lambda_1^*}NV_1N = QQ' - \frac{1}{\lambda_1^*}QQ'V_1QQ' =$

$$= Q(I - \frac{1}{\lambda_1^*}Q'V_1Q)Q' = Q(I - M_1^*)Q' = QM_2^*Q'.$$

2. $\frac{1}{\lambda_1^*}NV_1N = \frac{1}{\lambda_1^*}QQ'V_1QQ' = QM_1^*Q'.$

Mit $\delta_1 = 0$ und $\delta_2 = 1$ ist somit $\frac{1}{f} \cdot y'QM_2^*Q'y$ ein optimaler

Schätzer für σ_2^2 und mit $\delta_1 = \lambda_1^*$ und $\delta_2 = 1$ ist $\frac{\lambda_1^*}{q} y'QM_1^*Q'y$ ein optimaler Schätzer für $\lambda_1^*\sigma_1^2 + \sigma_2^2$.

Nach [5;S.63] gelten im BIB-Design $n = r \cdot t = b \cdot k$ und $\lambda \cdot (t-1) = r(k-1)$, woraus mit $b = t$ folgt $\frac{(t-1)}{r} \cdot \lambda t = n - t$. Da $f = q - rg\, M_1^*$ und $\frac{q}{\lambda_1^*} = q - rgM_2^*$ gelten, erhält man folgendes

2.3.7. Resultat:

Die optimalen Schätzer von σ_2^2 bzw. $\lambda_1^*\sigma_1^2 + \sigma_2^2$ erhält man durch die mittels der Methode der kleinsten Quadrate geschätzten Störvarianzen in den Modellen

$$Q'y = C_1^*\alpha_2 + e \text{ bzw. } Q'y = C_2^*\alpha_1 + e, \text{ wobei}$$
$$C_i^*C_i^{*\prime} = M_i^* \text{ für } i = 1,2 \text{ gelten.}$$

In [4;S.799] wird ausgeführt, daß kleinste Quadrate Schätzer in einem Regressionsmodell empfindlich gegenüber Ausreißern im Modellvektor sind, und es wird auf S.815 ein iteratives Verfahren zur Schätzung des Parametervektors der festen Effekte und der Störvarianz angegeben, dessen Lösung zu einem Schätzer der Störvarianz führt, welcher robust gegenüber Ausreißern im Modellvektor ist.

Wendet man dieses Verfahren auf die in 2.3.7. angegebenen Modelle an, so erhält man robuste Schätzer für σ_2^2 und $\lambda_1^*\sigma_1^2 + \sigma_2^2$.

2.4. Untersuchungen bei hierarchischen Designs

Bei einer einfach hierarchischen Klassifikation eines reinen Varianzkomponentenmodells ist die j-te Beobachtung des i-ten Individuums gegeben durch

$$y_{ij} = \mu + a_i + e_{ij} \text{ für } i = 1,2,\ldots,a \text{ und } j = 1,2,\ldots,n_i.$$

Faßt man die Zufallsvariablen a_i zu dem a-dimensionalen Zufallsvektor v_1 und die Zufallsvariablen e_{ij} zu dem n-dimensionalen Zufallsvektor v_2 mit $n := \sum_{i=1}^{a} n_i$ zusammen, so ist mit $U_1 := \sum_{i=1}^{a} \oplus 1_{n_i}$ und $U_2 := I_n$ der n-dimensionale Beobachtungsvektor y gegeben durch

$$y = 1_n \mu + U_1 v_1 + U_2 v_2 \text{ und sei im folgenden normalverteilt.}$$

Ferner sei angenommen, daß $E(v_1) = 0$, $E(v_2) = 0$, $E(v_1 v_1') = \sigma_1^2 I_a$, $E(v_2 v_2') = \sigma_2^2 I_n$ und $E(v_1 v_2') = 0$ gelten.

Damit ist $\mathrm{Cov}(y) = \sigma_1^2 V_1 + \sigma_2^2 V_2$ mit $V_1 := \sum_{i=1}^{a} \oplus J_{n_i}$ und $V_2 := I_n$.

Da $V_1 \underline{R}(X) = \underline{R}(n_1 1_{n_1}', n_2 1_{n_2}', \ldots, n_a 1_{n_a}')'$ ist, ist $\underline{R}(X) = \underline{R}(1_n)$ nur dann invariant bezüglich V_1, wenn $n_1 = n_2 = \ldots = n_a$ gilt. Wenn man also unbalanciertes Beobachtungsmaterial vorliegen hat, d.h. n_i nicht übereinstimmend ist für alle i, muß man die Voraussetzung von 2.2.9. untersuchen, um Bedingungen für optimale Schätzer angeben zu können. Zu diesem Zweck bestimmt man eine $(n \times q)$-Matrix Q mit $q = n-1$ so, daß $\underline{R}(Q) = \underline{N}(1_n') = \underline{R}(1_n)^{\perp}$ und $Q'Q = I_q$ gilt.

Die n-dimensionalen Spaltenvektoren c_k für $k = 1,2,\ldots,n-1$ einer solchen Matrix Q ergeben sich als orthonormalisierte Spaltenvektoren der (n×n)-Matrix $N := I_n - \frac{1}{n}J_n$ zu

$c_k := (\sqrt{(n-k)+(n-k)^2})^{-1} \quad (0,0,\ldots,-(n-k),1,1,\ldots,1)'$, wobei $-(n-k)$ die k-te Komponente des Vektors $(0,0,\ldots,-(n-k),1,1,\ldots,1)'$ ist.

2.4.1. Satz:

Bei einfach hierarchischer Klassifikation gibt es in der Klasse aller erwartungstreuen Schätzer für $\sum_{i=1}^{2} p_i \sigma_i^2$, welche Funktionen von Q'y sind, einen besten, wenn a = 2 ist.

Beweis: Nach 2.3.3. ist die lineare Hülle von $\{Q'V_1Q, I_a\}$ genau dann ein quadratischer Teilraum von $\mathfrak{C}$, wenn die Matrix $Q'V_1Q$ nicht mehr als zwei verschiedene Eigenwerte hat.

Die Spektraldarstellung von $Q'V_1Q$ ist gegeben durch

$$Q'V_1Q = \frac{2n_1 \cdot n_2}{n} \cdot \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1n_1} & 0 & \cdots & 0 \\ \cdot & q_{21} & \cdots & q_{2n_1} & 0 & \cdots & 0 \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdots & q_{n_1 n_1} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

mit $q_{ij} = \frac{n \cdot n_2}{n_1 \cdot \sqrt{[(n-i)+(n-i)^2]\cdot[(n-j)+(n-j)^2]}}$ für $i,j = 1,2,\ldots,n_1$.

Dabei hat der Eigenwert 0 die Vielfachheit n_1 und der Eigenwert $\frac{2n_1 \cdot n_2}{n}$ die Vielfachheit $n_2 - 1$.

Mit vollständiger Induktion zeigt man, daß $\sum_{l=1}^{n_1} \frac{1}{(n-1)+(n-1)^2} =$

$= \frac{n_1}{n(n-n_1)}$ gilt.

Damit kann man nachweisen, daß das Produkt der i-ten Zeile mit der j-ten Spalte obiger Matrix das (i,j)-te Element derselben Matrix ergibt, d.h. die Matrix ist idempotent.

2.4.2. Beispiel für a = 2:

$$\begin{pmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{21} \\ e_{22} \end{pmatrix} = 1_5 \cdot \mu + U_1 v_1 + U_2 v_2. \text{ Mit}$$

$$Q = \begin{pmatrix} \frac{-4}{\sqrt{20}} & 0 & 0 & 0 \\ \frac{1}{\sqrt{20}} & \frac{-3}{\sqrt{12}} & 0 & 0 \\ \frac{1}{\sqrt{20}} & \frac{1}{\sqrt{12}} & \frac{-2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{20}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{20}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{pmatrix} \text{ und}$$

$$Q'V_1Q = \frac{12}{5}\begin{pmatrix} \frac{1}{6} & \frac{10}{3\sqrt{20\cdot 12}} & \frac{10}{3\sqrt{20\cdot 6}} & 0 \\ \cdot & \frac{5}{18} & \frac{10}{3\sqrt{12\cdot 6}} & 0 \\ \cdot & \cdot & \frac{5}{9} & 0 \\ \cdot & \cdot & \cdot & 0 \end{pmatrix}$$

gelten $(Q'V_1Q)^2 = \frac{12}{5}\, Q'V_1Q$ und $(Q'V_1Q+I_4)^2 = \frac{22}{5}\, Q'V_1Q + I_4$, d.h. die lineare Hülle von $\{Q'V_1Q, I_4\}$ ist ein quadratischer Teilraum der Menge aller reellen und symmetrischen (4×4)-Matrizen.

2.4.3. Beispiel für a = 3:

$$\begin{pmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{31} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ a \\ a_1 \\ a_3 \end{pmatrix} + \begin{pmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{31} \end{pmatrix} = 1_4\cdot\mu + U_1v_1 + U_2v_2. \text{ Mit}$$

$$Q'V_1Q = \begin{pmatrix} \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ ist } (Q'V_1Q+I_4)^2 = \begin{pmatrix} \frac{11}{4} & \frac{7}{2\sqrt{2}} & 0 \\ \frac{7}{2\sqrt{2}} & \frac{9}{2} & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Annahme: Es gäbe ξ_1 und ξ_2 mit $(Q'V_1Q+I_4)^2 = \xi_1 Q'V_1Q+\xi_2 I_4$, dann ergibt ein Vergleich der ersten beiden Elemente der Hauptdiagonalen $\xi_1 = \frac{7}{2}$ und $\xi_2 = 1$.
Da aber $4 \neq \frac{7}{2}$ 1+1 ist, ist $(Q'V_1Q+I_4)^2$ nicht als Linearkombination von $Q'V_1Q$ und I_4 darstellbar, d.h. nach 2.2.4. ist die lineare Hülle von $\{Q'V_1Q, I_4\}$ kein quadratischer Teilraum der Menge aller reellen und symmetrischen (3×3)-Matrizen.

Es wird nun versucht, bei einer einfach hierarchischen Klassifikation 2.3.1. anzuwenden. Zu diesem Zweck ermittelt man die Spektraldarstellung der Kovarianzmatrix des Beobachtungsvektors einer einfach hierarchischen Klassifikation:

Zuerst bestimmt man die Spektraldarstellung von $V_1 = \overset{a}{\underset{i=1}{\oplus}} J_{n_i}$ und erhält $V_1 = 0\,(I - \sum_{i=1}^{a} M_i) + \sum_{i=1}^{a} n_i M_i$, wobei M_i diejenige unterteilte (n×n)-Matrix ist, bei der an der i-ten Stelle der Hauptdiagonale $\frac{1}{n_i} J_{n_i}$ und außerhalb Nullmatrizen stehen. Dabei hat der Eigenwert 0 die Vielfachheit n-a und die Eigenwerte n_i für i = 1,2,...,a haben je die Vielfachheit 1. Da $Cov(y) = \sigma_1^2 V_1 + \sigma_2^2 I_n$ ist, ergibt sich die

2.4.4. Spektraldarstellung der Kovarianzmatrix einer einfach hierarchischen Klassifikation:

$Cov(y) = \sum_{i=1}^{a} (n_i\sigma_1^2+\sigma_2^2) M_i + \sigma_2^2 (I_n - \sum_{i=1}^{a} M_i)$, wobei die Vielfachheit von $\lambda_i := n_i\sigma_1^2 + \sigma_2^2$ gleich 1 ist für i = 1,2,...,a und die Vielfachheit von σ_2^2 ist n-a.

Die zu den Eigenwerten λ_i gehörigen Eigenvektoren lauten

$C_i := (0,0,\ldots,0, \underbrace{\frac{1}{\sqrt{n_i}}, \frac{1}{\sqrt{n_i}}, \ldots, \frac{1}{\sqrt{n_i}}}_{n_i}, 0,0,\ldots,0)'$, wobei der

Teilvektor $\frac{1}{\sqrt{n_i}} \cdot 1_{n_i}$ an der i-ten Stelle für i = 1,2,...,a steht

Da die Eigenwerte von Cov(y) nur schätzbar sind, wenn ihre Vie

fachheit größer als 1 ist, ergibt sich die

2.4.5. Folgerung: Der beste quadratische erwartungstreue Schätzer von σ_2^2 ist gegeben durch

$\hat{\sigma}_2^2 := \frac{y'[I-Z(Z'Z)^{-1}Z']y}{n-rgZ}$, wobei $Z := \bigoplus_{i=1}^{a} \frac{1}{\sqrt{n_i}} \cdot 1_{n_i}$ (unter Weglas-

sung der Einsenspalte 1_n).

Da $Z'Z = I_a$ ist, gilt $\hat{\sigma}_2^2 = \frac{y'[I-ZZ']y}{n-a} = \frac{y'[I-P_1]y}{n-a}$,

d.h. $\hat{\sigma}_2^2$ ist der in 1.4. angegebene HENDERSON-Schätzer.

Bei einer zweifach hierarchischen Klassifikation ist die Beobachtung y_{ijk} gegeben durch

$$y_{ijk} = \mu + a_i + b_{ij} + e_{ijk} \text{ für } i = 1,2,\ldots,a;\ j = 1,2,\ldots$$

und $k = 1,2,\ldots,n_{ij}$. (Man betrachtet a "Stufen" eines "Faktors innerhalb des "Faktors A" tritt der "Faktor B" in b_i "Stufen" oder "Unterklassen" auf und in jeder "Unterklasse" gibt es n_{ij} Beobachtungen).

Im folgenden sei $n_{ij} = n_i$ für alle j.

Nach 1.1.1. ist die Matrixform des Modells gegeben durch

$y = 1\cdot\mu + \sum_{i=1}^{3} U_i v_i$ mit $Cov(y) = \sum_{i=1}^{3} \sigma_i^2 V_i$, wobei $V_1 := \overset{a}{\underset{i=1}{\Sigma}}{}^{\oplus} J_{b_i\cdot n_i}$,

$V_2 := \sum_{i=1}^{a} (I_{b_i} \otimes J_{n_i})$ und $V_3 := I_n$ ist.

Um die Spektraldarstellung von Cov(y) anzugeben, werden folgende Bezeichnungen eingeführt:

$R_i := \frac{1}{b_i\cdot n_i} J_{b_i\cdot n_i}$, $S_i := \frac{1}{n_i}(I_{b_i} \otimes J_{n_i}) - \frac{1}{b_i\cdot n_i} J_{b_i\cdot n_i}$ und

$T_i := I_{b_i\cdot n_i} - \frac{1}{n_i}(I_{b_i} \otimes J_{n_i})$ für $i = 1,2,\ldots,a$.

Ferner sei $\lambda_i := b_i\cdot n_i\sigma_1^2 + n_i\cdot\sigma_1^2 + \sigma_2^2$, $\nu_i := n_i\sigma_2^2 + \sigma_3^2$ und

$\mu_i := \sigma_3^2$ für $i = 1,2,\ldots a$.

2.4.6. Satz:

Die Spektraldarstellung von Cov(y) bei obigen Besetzungszahlen ist gegeben durch

$$Cov(y) = \overset{a}{\underset{i=1}{\Sigma}}{}^{\oplus} \lambda_i R_i + \overset{a}{\underset{i=1}{\Sigma}}{}^{\oplus} \nu_i S_i + \overset{a}{\underset{i=1}{\Sigma}}{}^{\oplus} \mu_i T_i.$$

Dabei haben die Eigenwerte λ_i, ν_i bzw. μ_i von Cov(y) die Vielfachheit 1, b_i-1 bzw. $n - \sum_{i=1}^{a} b_i$.

Beweis: Es gelten $R_i + S_i + T_i = I_{b_i\cdot n_i}$ für $i = 1,2,\ldots,a$,

$R_i\cdot S_i = R_i - R_i = 0$ für $i = 1,2,\ldots a$,

$S_i\cdot T_i = S_i - S_i = 0$ für $i = 1,2,\ldots a$ und

$R_i\cdot T_i = R_i(I_{b_i\cdot n_i} - R_i - S_i) = 0$ für $i = 1,2,\ldots a$.

Ferner sind $R_i^2 = R_i$, $S_i^2 = S_i$ und $T_i^2 = T_i$ für $i = 1,2,\ldots a$.

Damit lassen sich nach 2.3.1'. Schätzer für ν_i und σ_3^2 angeben.

2.5. Schätzer von Varianzkomponenten und ihr GAUSS'sches Risiko

Unter dem GAUSS'schen Risiko (mean square error) G des Schätzers $y'Ay$ für $\sum_{i=1}^{l} p_i\sigma_i^2$ versteht man den Ausdruck $G := E(y'Ay - \sum_{i=1}^{l} p_i\sigma_i^2)^2$.

2.5.1. Satz:

Sei y ein normalverteilter n-dimensionaler Zufallsvektor mit der Darstellung $y = X\beta + \sum_{i=1}^{l} U_i v_i$ und den Annahmen von 1.1.1., Ist $y'Ay$ ein Schätzer für $\sum_{i=1}^{l} p_i\sigma_i^2$, so ist das GAUSS'sche Risiko G von $y'Ay$ gegeben durch

$$G = 2\mathrm{tr}(A\,\mathrm{Cov}(y))^2 + 4\beta'X'A\,\mathrm{Cov}(y)\,AX\beta + \\ + [\mathrm{tr}(A\,\mathrm{Cov}(y)) + \beta'X'AX\beta - \sum_{i=1}^{l} p_i\sigma_i^2]^2.$$

Beweis:

$$E[y'Ay - \sum_{i=1}^{l} p_i\sigma_i^2]^2 =$$

$$= E[(y'Ay - E(y'Ay)) + (E(y'Ay) - \sum_{i=1}^{l} p_i\sigma_i^2)]^2 =$$

$$= \mathrm{var}(y'Ay) + [E(y'Ay) - \sum_{i=1}^{l} p_i\sigma_i^2]^2, \text{ wobei nach [9;S.57]}$$

$$\mathrm{var}(y'Ay) = 2\mathrm{tr}(A\,\mathrm{Cov}(y))^2 + 4\beta'X'A\,\mathrm{Cov}(y)\,AX\beta \text{ gilt.}$$

2.5.2. Folgerung:

Ist $y'Ay$ ein Schätzer für $\sum_{i=1}^{l} p_i\sigma_i^2$ mit $AX = 0$, so ist das GAUSS'sche Risiko G^* von $y'Ay$ gegeben durch

$$G^* = 2\mathrm{tr}(A\,\mathrm{Cov}(y))^2 + \left[\mathrm{tr}(A\,\mathrm{Cov}(y)) - \sum_{i=1}^{l} p_i\sigma_i^2\right]^2.$$

Zur Minimierung von G^* unter den Nebenbedingungen $AX = 0$ und $A = A'$ verwendet man den

2.5.3. Hilfssatz:

Bezeichnet man mit P den Orthogonalprojektor in $\underline{R}(X)$ bezüglich des Skalarproduktes $(a,b) = a'Cov(y)^{-1}b$, so ist notwendig und hinreichend für $AX = 0$ und $A = A'$ die Darstellung $A = Cov(y)^{-1}(I-P)\cdot S\cdot Cov(y)^{-1}(I-P)$ mit symmetrischer $(n\times n)$-Matrix S.

Beweis: Sei $AX = 0$ und $A = A'$, dann ist mit $S = Cov(y)\cdot A\cdot Cov(y)$ $Cov(y)^{-1}(I-P)\cdot Cov(y)\cdot A\cdot Cov(y)\cdot Cov(y)^{-1}(I-P) = A$.

Ist umgekehrt $A = Cov(y)^{-1}(I-P)\cdot S\cdot Cov(y)^{-1}(I-P)$, so ist $AX = 0$ und $A = A'$.

Sei F eine partiell differenzierbare Funktion der Variablen s_{ij}, die in der $(n\times n)$-Matrix S angeordnet sind. Bildet man die partiellen Ableitungen $\frac{\partial F}{\partial s_{ij}}$ für $i,j = 1,2,\ldots,n$ und ordnet sie in derselben Weise wie die s_{ij} zu einer $(n\times n)$-Matrix, die man mit $\frac{\partial F}{\partial S}$ bezeichnet, so gelten folgende Ableitungsregeln, welche sich als Spezialfälle der in [8;S.72] angegebenen Regeln ergeben.
Sind S und B symmetrische Matrizen, so gelten

(i) $\frac{\partial tr(SB)^2}{\partial S} = 4BSB - 2\ diag(BSB)$

(ii) $\frac{\partial tr(SB)}{\partial S} = 2B - diag\ B$, wobei diag(A) diejenige Diagonalmatrix mit derselben Hauptdiagonalen wie A ist.

Es wird nun folgendes Resultat von RAO [6;S.451] bewiesen:

2.5.4. Satz:

Unter allen symmetrischen (n×n)-Matrizen mit AX = 0 minimiert

$A^* := \frac{\sum_{i=1}^{l} p_i \sigma_i^2}{n-rgX+2} \cdot Cov(y)^{-1}(I-P)$ das in 2.5.2. angegebene GAUSS'sche Risiko G^*.

Beweis: Mit $A = Cov(y)^{-1}(I-P) \cdot S \cdot Cov(y)^{-1}(I-P)$ ist

$$G^* = 2tr(S \cdot Cov(y)^{-1}(I-P) \cdot S \cdot Cov(y)^{-1}(I-P)) +$$

$$+ [tr(S \cdot Cov(y)^{-1}(I-P)) - \sum_{i=1}^{l} p_i \sigma_i^2]^2 \text{ und}$$

$$\frac{\partial G^*}{\partial S} = 2 \cdot [4A - 2 \operatorname{diag} A] + 2[tr(A\, Cov(y)) - \sum_{i=1}^{l} p_i \sigma_i^2] \cdot$$

$$\cdot [2\, Cov(y)^{-1}(I-P) - \operatorname{diag}((Cov(y)^{-1}(I-P))]$$

$A^* = \frac{\sum_{i=1}^{l} p_i \sigma_i^2}{n-rgX+2} Cov(y)^{-1}(I-P)$ in $\frac{\partial G}{\partial S}$ eingesetzt ergibt

wegen $tr(Cov(y)^{-1}(I-P)\, Cov(y)) = tr(I-P) = rg(I-P) = n - rgX$ die Bedingung $\frac{\partial G^*}{\partial S} = 0$.

Die Spektraldarstellung der Varianz-Kovarianzmatrix einer einfach hierarchischen Klassifikation lautet nach

2.4.4.:

$$Cov(y) = \sum_{i=1}^{a} (n_i \sigma_1^2 + \sigma_2^2) \cdot M_i + \sigma_2^2 (I_n - \sum_{i=1}^{a} M_i), \text{ mit } \sum_{i=1}^{a} M_i = \bigoplus_{i=1}^{a} \frac{1}{n_i} J_{n_i}$$

Damit ist $\mathrm{Cov}(y)^{-1} = \sum_{i=1}^{a} \frac{1}{n_i\sigma_i^2+\sigma_2^2} \cdot M_i + \frac{1}{\sigma_2^2}(I_n - \sum_{i=1}^{a} M_i)$.

Da $(I_n - \sum_{i=1}^{a} M_i)1_n = 0$, ist

$\mathrm{Cov}(y)^{-1} 1_n = \sum_{i=1}^{a} \frac{1}{n_i\sigma_1^2+\sigma_2^2} M_i 1_n$ und $1_n'\mathrm{Cov}(y)^{-1}1_n = \sum_{i=1}^{a} \frac{n_i}{n_i\sigma_1^2+\sigma_2^2}$

Somit ist diejenige Matrix A^* mit $A^* 1_n = 0$, die den Schätzer $y'A^* y$ von σ_2^2 mit kleinstem GAUSS'schen Risiko liefert, gegeben durch

$$A^* = \frac{\sigma_2^2}{n+1}\Big[\sum_{i=1}^{a} \frac{1}{n_i\sigma_1^2+\sigma_2^2} M_i + \frac{1}{\sigma_2^2}(I_n - \sum_{i=1}^{a} M_i) -$$

$$- \frac{1}{\sum_{i=1}^{a} \frac{n_i}{n_i\sigma_1^2+\sigma_2^2}} \Big(\sum_{i=1}^{a} \frac{1}{n_i\sigma_1^2+\sigma_2^2} m_i \Big)\Big(\sum_{i=1}^{a} \frac{1}{n_i\sigma_1^2+\sigma_2^2} m_i' \Big) \Big]$$

wobei $m_i' := (0,\ldots,0,1,\ldots,1,0,\ldots,0)$ ein $a \cdot n_i$-Vektor mit n_i Einsen ist.

Wählt man $n_i = b$ für $i = 1,2,\ldots,a$, so erhält man

$$A^* = \frac{\sigma_2^2}{n+1}\Big[\frac{1}{b\sigma_1^2+\sigma_2^2} (I_a \otimes \tfrac{1}{b}J_b) + \frac{1}{\sigma_2^2}(I-(I_a \otimes \tfrac{1}{b}J_b)) - \frac{1}{b\sigma_1^2+\sigma_2^2} \tfrac{1}{n} J_n \Big]$$

$$= \frac{\sigma_2^2}{n+1}\Big[\frac{1}{b\sigma_1^2+\sigma_2^2}((I_a \otimes \tfrac{1}{b}J_b) - \tfrac{1}{n} J_n)\Big] + \frac{1}{n+1}\Big[I-(I_a \otimes \tfrac{1}{b}J_b)\Big].$$

2.5.5. Folgerung:

Dann und nur dann, wenn $a = 1$ und $b = n$ ist, ist A^* unabhängig von den unbekannten Varianzkomponenten.

Genau in diesem Fall ist $I_a \otimes \frac{1}{b}J_b - \frac{1}{n}\cdot J_n = 0$.

Der damit erhaltene Schätzer $y' \frac{1}{n+1}(I-\frac{1}{n}J_n)y$ für σ_2^2 ist ein Vielfaches des HENDERSON-Schätzers aus 1.4.

Jetzt wird $A^{\blacksquare}$ in dem Spezialfall untersucht, daß die Beobachtung y_{ij} eines quantitativen Merkmals gegeben ist durch

$$y_{ij} = \mu + a_i + b_j + e_{ij} \text{ für } i = 1,2,\ldots,a \text{ und } j = 1,2,\ldots,b.$$

("Kreuzklassifikation")

a_i, b_j und e_{ij} seien unkorrelierte, normalverteilte Zufallsvariablen mit dem Erwartungswert Null und den Varianzen σ_1^2, σ_2^2 und σ_3^2.

Für die Varianz-Kovarianzmatrix Cov(y) des Beobachtungsvektors y gelte

$$\text{Cov}(y) = \sigma_1^2(I_a \otimes J_b) + \sigma_2^2(J_a \otimes I_b) + \sigma_3^2 I_n \text{ mit } n := a\cdot b.$$

Mit den Abkürzungen $\lambda_1 := b\sigma_1^2 + a\sigma_2^2 + \sigma_3^2$, $\lambda_2 := a\sigma_2^2 + \sigma_3^2$,

$$\lambda_3 := b\sigma_1^2 + \sigma_3^2, \qquad \lambda_4 := \sigma_3^2,$$

$$M_1 := \frac{1}{n}J_n,\ M_2 := \frac{1}{a}J_a \otimes I_b - \frac{1}{n}J_n,$$

$$M_3 := I_a \otimes \frac{1}{b}J_b - \frac{1}{n}J_n \text{ und}$$

$$M_4 := I_n - \frac{1}{a}J_a \otimes I_b - I_a \otimes \frac{1}{b}J_b + \frac{1}{n}J_n.$$

lautet die Spektraldarstellung der Varianz-Kovariantmatrix des Beobachtungsvektors y

$$\text{Cov}(y) = \sum_{i=1}^{4} \lambda_i M_i.$$

Wegen $(I_a \otimes \frac{1}{b}J_b)\cdot 1_n = 1_n$ und $(\frac{1}{a}J_a \otimes I_b)\cdot 1_n = 1_n$

gelten $M_1 1_n = 1_n$, $M_2 1_n = 0$, $M_3 1_n = 0$ und $M_4 1_n = 0$.

Damit ist $\mathrm{Cov}(y)^{-1}1_n = \frac{1}{\lambda_1}1_n$, $1_n' \mathrm{Cov}(y)^{-1}1_n = \frac{n}{\lambda_1}$ und nach 2.5.4.

$$A^* = \frac{\sum_{i=1}^{3} p_i\sigma_i^2}{n+1}\left[\mathrm{Cov}(y)^{-1} - \mathrm{Cov}(y)^{-1}1_n(1_n'\mathrm{Cov}(y)^{-1}1_n)^{-1}1_n'\mathrm{Cov}(y)^{-1}\right]$$

$$= \frac{\sum_{i=1}^{3} p_i\sigma_i^2}{n+1}\left[\frac{1}{\lambda_2}M_2 + \frac{1}{\lambda_3}M_3 + \frac{1}{\lambda_4}M_4\right].$$

Schätzt man λ_2 bzw. λ_3 mittels $y'A^*y$, so ergibt sich für $a = 1$ und $b = n$ bzw. $a = n$ und $b = 1$

$A^* = \frac{1}{n+1}(I_n - \frac{1}{n}J_n)$. Nur in diesen Spezialfällen ist A^* unabhängig von den unbekannten Varianzkomponenten.

Da die Minimierung von G^* nur bei sehr eingeschränkten Datenstrukturen zum Erfolg geführt hat, sei jetzt das Varianzkomponentenmodell einer einfach hierarchischen Klassifikation mit balanziertem Beobachtungsmaterial zugrunde gelegt. Es sei also $n_i = b$ für alle $i = 1,2,\ldots,a$, ferner $U_1 := I_a \otimes 1_b$ und $V_1 := I_a \otimes J_b$. Der n-dimensionale Zufallsvektor y sei normalverteilt mit $E(y) = 1_n\cdot\mu$ und $\mathrm{Cov}(y) = \sigma_1^2V_1 + \sigma_2^2I_n$.

Nach 1.4. ist HENDERSON's Schätzer nach Methode III für σ_2^2 gegeben durch $y'N_2y = \hat{\sigma}_2^2 \mathrm{tr}N_2$ mit $N_2 = I_n - P_1$, wobei $P_1 := \frac{1}{b}(I_a \otimes J_b)$ und $\mathrm{tr}\,N_2 = n-a$ ist.

Damit ist $\hat{\sigma}_2^2 = y'\frac{N_2}{n-a}y$ ein erwartungstreuer Schätzer für σ_2^2 und da das GAUSS-sche Risiko in diesem Fall gleich der Varianz von $y'\frac{N_2}{n-a}y$ ist, gilt

$$E\left[y'\frac{N_2}{n-a}y-\sigma_2^2\right]^2 = 2\mathrm{tr}\left(\frac{N_2}{n-a}(\sigma_1^2V_1+\sigma_2^2I)\right)^2 = \frac{2\sigma_2^4}{(n-a)}\mathrm{tr}N_2 = \frac{2\sigma_2^4}{n-a},$$

da $N_2V_1 = (I_a \otimes J_b) - \frac{1}{b}\cdot b\cdot(I_a \otimes J_b) = 0$.

2.5.6. Satz:

Unter allen Schätzern $y'\frac{N_2}{c}y$ von σ_2^2 hat unter den oben genannten Annahmen der Schätzer mit $c = n-a+2$ minimales GAUSS-sches Risiko.

Beweis: Ist $E\left[y'\frac{N_2}{c}y-\sigma_2^2\right]^2 = 2\mathrm{tr}\left(\frac{\sigma_2^2}{c}\cdot N_2\right)^2 + \left[\mathrm{tr}\left(\frac{\sigma_2^2}{c}\cdot N_2\right)-\sigma_2^2\right]^2$ minimal,

dann ist $\frac{dE\left[y'\frac{N_2}{c}y-\sigma_2^2\right]^2}{dc} = -\frac{4(n-a)\sigma_2^4}{c^3} - \frac{2(n-a)^2\sigma_2^4}{c^3} + \frac{2(n-a)\sigma_2^4}{c^2} = 0$

und somit $c = n-a+2$.

Ferner ist $\frac{d^2E\left[y'\frac{N_2}{c}y-\sigma_2^2\right]^2}{dc^2}$ an der Stelle $c = n-a+2$ gegeben durch

$$\frac{d^2E\left[y'\frac{N_2}{n-a+2}y-\sigma_2^2\right]^2}{dc^2} = \frac{2(n-a)\cdot\sigma_2^4}{(n-a+2)^3} > 0, \text{ falls } n > a-2.$$

Nach 1.4. ist HENDERSON's Schätzer nach Methode III für σ_1^2 gegeben durch

$y'N_1y = \hat{\sigma}_1^2 \, tr(N_1V_1) + \hat{\sigma}_2^2 trN_1$, wobei $N_1 = P_1 - P_o$ mit $P_o = \frac{1}{n}J_n$ und $P_1 = \frac{1}{b}(I_a \otimes J_b)$.

Da $N_1V_1 = \left[\frac{1}{b}(I_a \otimes J_b) - \frac{1}{n}J_n\right](I_a \otimes J_b) = (I_a \otimes J_b) - \frac{b}{n}J_n = b \cdot N_1$,

$tr(N_1V_1) = n-b$ und $trN_1 = a-1$,

ist $\hat{\sigma}_1^2 = y'\left[\frac{N_1}{n-b} - \frac{N_2}{\frac{(n-b)(n-a)}{a-1}}\right]y$ ein

erwartungstreuer Schätzer für σ_1^2.

Das GAUSS-sche Risiko des Schätzers $y'\left[\frac{N_1}{f} - \frac{N_2}{g}\right]y$ für σ_1^2 ist gegeben durch

$$G(f,g) := E\left[y'\left(\frac{N_1}{f} - \frac{N_2}{g}\right)y - \sigma_1^2\right]^2 =$$

$$= \frac{2(b\sigma_1^2+\sigma_2^2)^2}{f^2}(a-1) + \frac{2\sigma_2^4}{g^2}(n-a) +$$

$$+ \left[\frac{a-1}{f}(b\sigma_1^2+\sigma_2^2) - \frac{(n-a)}{g}\sigma_2^2 - \sigma_1^2\right]^2,$$

da $N_1N_2 = 0$ und $N_2V_1 = 0$ gelten.

$$\frac{\partial G(f,g)}{\partial f} = \frac{-4(b\sigma_1^2+\sigma_2^2)^2 \cdot (a-1)}{f^3} - 2\cdot\left[\frac{(a-1)\cdot(b\sigma_1^2+\sigma_2^2)}{f} - \frac{(n-a)}{g}\sigma_2^2 - \sigma_1^2\right] \cdot$$

$$\cdot \frac{(a-1)\cdot(b\sigma_1^2+\sigma_2^2)}{f^2} = 0$$

ergibt $f_1 := b(a+1)$ und $g_1 := b(n-a)$.

$$\frac{\partial G(f,g)}{\partial g} = \frac{-4\sigma_2^4(n-a)}{g^3} + 2\left[\frac{(a-1)\cdot(b\sigma_1^2+\sigma_2^2)}{f} - \frac{(n-a)}{g}\sigma_2^2 - \sigma_1^2\right]\cdot\frac{(n-a)}{g^2}\sigma_2^2 =$$

ergibt $f_2 := n-b$ und $g_2 := b(n-a+2)$.

Zur Berechnung der Funktionaldeterminante der partiellen Ableitungen

$\frac{\partial G(f,g)}{\partial f}$ und $\frac{\partial G(f,g)}{\partial g}$ benötigt man

$$\frac{\partial^2 G(f,g)}{\partial f^2} = \frac{2(b\sigma_1^2+\sigma_2^2)(a-1)}{f^3}\cdot\left[\frac{6(b\sigma_1^2+\sigma_2^2)}{f} + \frac{3(a-1)\cdot(6\sigma_1^2+\sigma_2^2)}{f} -\right.$$

$$\left. - \frac{2(n-a)\sigma_2^2}{g} - 2\sigma_1^2\right],$$

$$\frac{\partial^2 G(f,g)}{\partial g^2} = \frac{2(n-a)\sigma_2^4}{g^4}\cdot(6+3(n-a)) + \frac{4(n-a)}{g^3}\left[\sigma_1^2\sigma_2^2 - \frac{\sigma_2^2(a-1)\cdot(b\sigma_1^2+\sigma_2^2)}{f}\right] u$$

$$\frac{\partial^2 G(f,g)}{\partial f \partial g} = -\frac{2(n-a)\cdot(a-1)\cdot(b\sigma_1^2+\sigma_2^2)\cdot\sigma_2^2}{f^2\cdot g^2}.$$

Ferner ist $\frac{\partial^2 G(f_1,g_1)}{\partial f^2}\cdot\frac{\partial^2 G(f_1g_1)}{\partial g^2} - \left[\frac{\partial^2 G(f_1g_1)}{\partial f \partial g}\right]^2 > 0$

und $\frac{\partial^2 G(f_1,g_1)}{\partial f^2} = \frac{2(b\sigma_1^2+\sigma_2^2)^2\cdot(a-1)}{b^4\cdot(a+1)^3} > 0.$

Damit hat das GAUSS-sche Risiko $G(f,g)$ an der Stelle (f_1,g_1) ein Minimum im engeren Sinne und es gilt

$$G(f_1,g_1) = \frac{2(b\sigma_1^2+\sigma_2^2)^2}{b^2(a+1)} + \frac{2\sigma_2^4}{b^2(n-a)}.$$

Der Wert des GAUSS-schen Risikos an der Stelle (f_2,g_2) ist gegeben durch

$$G(f_2,g_2) = \frac{2(b\sigma_1^2+\sigma_2^2)^2\cdot(a-1)}{(n-b)^2\cdot} + \frac{2\sigma_2^4}{b^2\cdot(n-a+2)}$$

und für $f_3 := n-b$ und $g_3 := \frac{(n-b)\cdot(n-a)}{a-1} = b(n-a)$ ergibt sich

$$G(f_3,g_3) = \frac{2(b\sigma_1^2+\sigma_2^2)^2\cdot(a-1)}{(n-b)^2} + \frac{2\sigma_2^4(a-1)^2}{(n-b)^2\cdot(n-a)}.$$

Ein Vergleich der GAUSS-schen Risiken für (f_1,g_1), (f_2,g_2) und (f_3,g_3) ergibt $G(f_1,g_1) < G(f_2,g_2) < G(f_3,g_3) = \text{var}\,\hat{\sigma}_1^2$.

Als Resultat ergibt sich somit

2.5.7. Satz:

Unter allen Schätzern $y'\left[\frac{N_1}{f} - \frac{N_2}{g}\right]y$ von σ_1^2 hat unter den genannten Annahmen der Schätzer mit $f = f_1 = b(a+1)$ und $g = g_1 = b(n-a)$ minimales GAUSS-sches Risiko.

Im folgenden sei ein Balanced Complete Block Design zugrunde gelegt mit den im Anschluß an 2.3.2. angegebenen Annahmen. Nach 1.4. ist HENDERSON's Schätzer nach Methode III für σ_2^2 gegeben durch $y'N_2y = \hat{\sigma}_2^2 \text{tr}N_2$, wobei $N_2 = I-P_1$ und P_1 der Orthogonalprojektor in $\underline{R}(X,U_1)$ ist.

Um P_1 anzugeben, seien die symmetrischen (b×b)-Matrizen

$$Q_1 := \begin{pmatrix} b+t-1 & t-1 & \dots & t-1 \\ t-1 & b+t-1 & \dots & t-1 \\ \vdots & \vdots & \ddots & \vdots \\ t-1 & t-1 & \dots & b+t-1 \end{pmatrix} \quad \text{und}$$

$$Q_2 := \begin{pmatrix} b-1 & -1 & \dots & -1 \\ -1 & b-1 & \dots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \dots & b-1 \end{pmatrix} \quad \text{eingeführt.}$$

Damit ist $P_i = \frac{1}{n}\cdot\left[(J_t \otimes Q_2) - (I_t \otimes Q_2) + (I_t \otimes Q_1)\right]$ und

$$trN_2 = trI_n - trP_1 = n - \frac{1}{n}\left[t\cdot b(b-1) - t\cdot b(b-1) + t\cdot b(b+t-1)\right] = n-b-t+1.$$

2.5.8. Satz:

Ist y normalverteilt mit $E(y) = X\beta$ und $Cov(y) = \sigma_1^2 V_1 + \sigma_2^2 I$, dann ist derjenige Schätzer von σ_2^2, der unter allen Schätzern $y'\frac{N_2}{c}y$ das kleinste GAUSS-sche Risiko hat, gegeben durch $c = n-b-t+3$.

Beweis: $E(y'\frac{N_2}{c}y - \sigma_2^2)^2 = \frac{2\sigma_2^4}{c^2}(n-b-t+1) + \left[\frac{\sigma_2^2(n-b-t+1-c)}{c}\right]^2 =: G(c)$

$$\frac{dG(c)}{dc} = \frac{-2\sigma_2^4}{c^3}(n-b-t+1)\cdot\left[2+(n-b-t+1-c)\right] = 0 \text{ liefert } c = n-b-t+3.$$

Ferner ist $\frac{d^2 G(c)}{dc^2} = \frac{2\sigma_2^4(n-b-t+1)}{c^4}\cdot\left[6+3(n-b-t+1)-2c\right]$ und

$$\frac{d^2G(n-b-t+3)}{dc^2} = \frac{2\sigma_2^4(n-b-t+1)}{(n-b-t+3)^3} > 0.$$

Nach 1.4. ist HENDERSON's Schätzer nach Methode III für σ_1^2 gegeben durch $y'N_1y = \hat{\sigma}_1^2 tr(N_1V_1) + \hat{\sigma}_2^2 trN_1$, wobei $N_1 = P_1 - P_o$, $P_o = \frac{1}{b}(I_t \otimes J_b)$, $tr(N_1) = b-1$, $V_1 = J_t \otimes I_b$, $N_1V_1 = t\cdot N_1$, $P_oV_1 = \frac{1}{b}J_n$. Damit ist $tr(N_1V_1) = t\cdot(b-1)$ und somit $\hat{\sigma}_1^2 = y'(\frac{N_1}{t(b-1)} - \frac{N_2}{(n-b-t+1)\cdot t})y$ ein erwartungstreuer Schätzer für σ_1^2.

Das GAUSS-sche Risiko des Schätzers $y'(\frac{N_1}{c} - \frac{N_2}{d})y$ für σ_1^2 ist gegeben durch

$$G(c,d) := E\left[y'(\frac{N_1}{c} - \frac{N_2}{d})y - \sigma_1^2\right]^2 =$$

$$= \frac{2(b-1)}{c^2}\cdot(t\sigma_1^2+\sigma_2^2)^2 + \frac{2\sigma_2^4(n-b-t+1)}{d^2} + \left[\frac{b-1}{c}(t\sigma_1^2+\sigma_2^2) - \frac{\sigma_2^2(n-b-t+1)}{d} - \sigma_1^2\right]^2$$

Aus $\frac{\partial G(c,d)}{\partial c} = \frac{-4(b-1)(t\sigma_1^2+\sigma_2^2)^2}{c^3} - 2\left[\frac{(b-1)}{c}(t\sigma_1^2+\sigma_2^2) - \frac{\sigma_2^2(n-b-t+1)}{d} - \sigma_1^2\right]\cdot$

$$\cdot\frac{(b-1)\cdot(t\sigma_1^2+\sigma_2^2)}{c^2} = 0$$

erhält man $c_1 = (b+1)\cdot t$ und $d_1 = (n-b-t+1)\cdot t$.

Aus $\frac{\partial G(c,d)}{\partial d} = \frac{-4\sigma_2^4(n-b-t+1)}{d^3} + 2\left[\frac{(b-1)}{c}(t\sigma_1^2+\sigma_2^2)-\frac{\sigma_2^2(n-b-t+1)}{d}-\sigma_1^2\right]\cdot$

$\cdot\frac{\sigma_2^2(n-b-t+1)}{d^2} = 0$ erhält man $c_2 = t(b-1)$ und $d_2 = (n-b-t+3)\cdot t$.

2.5.9. Satz:

Ist y normal verteilt mit $E(y) = X\beta$ und $Cov(y) = \sigma_1^2V_1+\sigma_2^2I$, dann ist derjenige Schätzer von σ_1^2, der unter allen Schätzern $y'(\frac{N_1}{c} - \frac{N_2}{d})y$ das kleinste GAUSS-sche Risiko hat, gegeben durch $c_1 = (b+1)\cdot t$ und $d_1 = (n-b-t+1)\cdot t$.

Beweis: $\frac{\partial^2G(c_1,d_1)}{\partial c^2}\cdot\frac{\partial^2G(c_1,d_1)}{\partial d^2} - \left[\frac{\partial^2G(c_1,d_1)}{\partial c\,\partial d}\right]^2 =$

$$\frac{2(b-1)\cdot(t\sigma_1^2+\sigma_2^2)^2}{(b+1)^3\cdot t^4}\cdot\left[\frac{6\sigma_2^4}{t^4\cdot(n-b-t+1)^2}\left(\frac{n-b-t+3}{n-b-t+1}-\frac{b-1}{b+1}\right)+\right.$$

$$\left.+\frac{8\sigma_1^2\sigma_2^2}{(b+1)\cdot(n-b-t+1)^2\cdot t^3}\right] > 0$$

$$\frac{\partial^2G(c_1,d_1)}{\partial c^2} = \frac{2(b-1)\cdot(t\sigma_1^2+\sigma_2^2)^2}{(b+1)^3\cdot t^4} > 0.$$

Ferner sind

$$G((b+1)\cdot t,\ (n-b-t+1)\cdot t) = \frac{2(t\sigma_1^2+\sigma_2^2)^2}{(b+1)\cdot t^2} + \frac{2\sigma_2^4}{(n-b-t+1)\cdot t^2},$$

$$G((b-1)\cdot t,\ (n-b-t+3)\cdot t) = \frac{2(t\sigma_1^2+\sigma_2^2)^2}{(b-1)\cdot t^2} + \frac{2\sigma_2^4}{(n-b-t+3)\cdot t^2}$$

und

$$G((b-1)\cdot t,\ (n-b-t+1)\cdot t) = \frac{2(t\sigma_1^2+\sigma_2^2)^2}{(b-1)\cdot t^2} + \frac{2\sigma_2^4}{(n-b-t+1)\cdot t^2}\ .$$

Somit ist, falls $b \geq 1$ und $t > 2$,

$$G((b+1)\cdot t,\ (n-b-t+1)\cdot t) < G((b-1)\cdot t,\ (n-b-t+3)\cdot t) <$$

$$G((b-1)\cdot t,\ (n-b-t+1)\cdot t).$$

3. Anwendung der Schätztheorie von Kovarianzmatrizen auf RCR-Modelle von P.A.V.B. SWAMY.

Die zeitliche Abhängigkeit eines quantitativen Merkmals y_{it} des i-ten Individuums zur Zeit t sei gegeben durch

$$y_{it} = \delta_o + \delta_1 x_{1it} + \delta_2 x_{2it} + \ldots + \delta_{m-1} x_{m-1,it} + e_{it}$$
$$\text{für } i = 1,2,\ldots,N \text{ und } t = 1,2,\ldots,T.$$

Dabei bedeuten $\delta_o, \delta_1, \ldots, \delta_{m-1}$ unbekannte Parameter, x_{kit} ist der Wert der k-ten unabhängigen Variablen des i-ten Individuums zur Zeit t und e_{it} sind für $i = 1,2,\ldots,N$, $t = 1,2,\ldots,T$ zufällige Störungen.

Dieser Ansatz geht von der Annahme aus, daß der m-dimensionale Parametervektor $\delta := (\delta_o, \delta_1, \ldots, \delta_{m-1})'$ fest und für alle N Individuen gleich ist. Dadurch wird weder dem unterschiedlichen Verhalten noch der räumlichen Heterogenität der einzelnen Individuen Rechnung getragen und die Anwendung dieses Ansatzes ist sehr beschränkt.

Realistischer ist die Darstellung

$$y_{it} = \delta_{oi} + \delta_{1i} x_{1it} + \delta_{2i} x_{2it} + \ldots + \delta_{m-1,i} x_{m-1,it} + e_{it}$$
$$\text{für } i = 1,2,\ldots,N \text{ und } t = 1,2,\ldots,T,$$

bei welcher der Parametervektor $\delta_i := (\delta_{oi}, \delta_{1i}, \ldots, \delta_{m-1,i})'$

für jedes der N Individuen fest ist, aber von Individuum zu Individuum variiert.

Damit kommen im neuen Modell N·m feste unbekannte Parameter vor, welche bei der Untersuchung von vielen Individuen mit Hilfe kurzer Zeitreihen die Schätzung derselben problematisch machen.

Einen Kompromiß zwischen beiden Ansätzen stellt die Annahme eines in [13;S.99ff] eingeführten

Random-Coefficient-Regression-Modells

dar. Man faßt die N Parametervektoren δ_i verschiedener Individuen als unabhängige Stichprobe derselben multivariaten Verteilung mit unbekanntem m-dimensionalem Erwartungswertvektor und unbekannter (m×m)-Kovarianzmatrix Σ auf. Damit ist die Anzahl der zu schätzenden festen Parameter reduziert und zugleich die Heterogenität zwischen den einzelnen Individuen berücksichtigt.

Der Schätztheorie legt man folgendes Modell zugrunde:

Die Beobachtung eines quantitativen Merkmals des i-ten Individuums zur Zeit t sei gegeben durch

$$y_{it} = \beta_o + b_{oi} + (\beta_1 + b_{1i})x_{1it} + \ldots + (\beta_{m-1} + b_{m-1,i})x_{m-1,it} + e_{it}$$

für $i = 1,2,\ldots,N$ und $t = 1,2,\ldots,T$.

Dabei bedeuten x_{kit} den Wert der k-ten unabhängigen Variablen des i-ten Individuums zur Zeit t; $\beta_o, \beta_1, \ldots, \beta_{m-1}$ sind die Elemente des unbekannten m-dimensionalen Parametervektors

$\beta := (\beta_0, \beta_1, \ldots, \beta_{m-1})'$, $b_{0i}, b_{1i}, \ldots, b_{m-1,i}$ sind die Elemente des m-dimensionalen Zufallsvektors $b_i := (b_{0i}, b_{1i}, \ldots, b_{m-1,i})'$ mit $E(b_i) = 0$ für $i = 1,2,\ldots,N$

$$\text{und } E(b_i b_j') = \begin{cases} \Sigma & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases}$$

wobei Σ die unbekannte (m×m)-Kovarianzmatrix der Zufallsvektoren b_i für $i = 1,2,\ldots,N$ ist.

Mit den Bezeichnungen $y_i := (y_{i1}, y_{i2}, \ldots, y_{iT})'$,

$$X_i := \begin{pmatrix} 1 & x_{1i1} & \cdots & x_{m-1,i1} \\ 1 & x_{1i2} & \cdots & x_{m-1,i2} \\ \vdots & \vdots & & \vdots \\ 1 & x_{1iT} & \cdots & x_{m-1,iT} \end{pmatrix}$$

und $e_i := (e_{i1}, e_{i2}, \ldots, e_{iT})'$ schreibt man für das i-te Individuum kurz $y_i = X_i(\beta + b_i) + e_i$, wobei man zusätzlich $E(e_i) = 0$, $E(e_i e_i') = \sigma^2_{N+i} \cdot I_T$ mit unbekanntem $\sigma^2_{N+i} > 0$ und $E(e_i e_j') = 0$ für $i \neq j$ annimmt. Ferner sei $E(e_i b_j') = 0$ für alle $i,j = 1,2,\ldots,N$.

Für den Gesamtbeobachtungsvektor $y := (y_1', y_2', \ldots, y_N')'$ der Länge $n := NT$ gilt mit den Bezeichnungen $X := (X_1', X_2', \ldots, X_N')'$ und $e := (e_1', e_2', \ldots, e_N')'$

$$y = X\beta + \bigoplus_{i=1}^{N} X_i b_i + e \text{ mit } E(y) = X\beta \text{ und}$$

$$\text{Cov}(y) = \bigoplus_{i=1}^{N} (X_i \Sigma X_i' + \sigma^2_{N+i} \cdot I_T).$$

Um nun die MINQUE-Theorie zur Schätzung von $\sum$ und σ^2_{N+i} für $i = 1,2,\ldots,N$ anzuwenden, schreibt man obiges Modell für kombinierte Zeitreihen- und Querschnittsdaten als Spezialfall des in 1.1.1. dargestellten allgemeinen linearen Modells:

Dazu setzt man $q = m$, $s = N$, $l = 2N$, $k_i = T$ für $i = N+1,\ldots,2N$ und definiert die unterteilte $(n\times m)$-Matrix $U_j := (0,0,\ldots,0, X_j',0,\ldots,0)'$ für $j = 1,2,\ldots,N$, wobei die $(T\times m)$-Matrix X_j an der j-ten Stelle steht.

Entsprechend sei die unterteilte $(n\times T)$-Matrix U_i definiert durch $U_i := (0,0,\ldots, I_T,0,\ldots,0)'$ für $i = N+1,N+2,\ldots,2N$, wobei die $(T\times T)$-Matrix I_T an der $(i-N)$-ten Stelle steht.

Damit lassen sich die Schätzgleichungen (1) und (2) aus 2.1.4. aufstellen.

Mit den sich daraus ergebenden Schätzungen $\widehat{\sum}, \widehat{\sigma}^2_{N+1}, \ldots, \widehat{\sigma}^2_{2N}$ berechnet man die MINQUE der Kovarianzmatrix des Beobachtungsvektors y durch

$$\widehat{Cov(y)} = \sum_{j=1}^{N} U_j \widehat{\sum} U_j' + \sum_{i=N+1}^{2N} \widehat{\sigma}^2_i U_i U_i'$$

und bestimmt den AITKEN-Schätzer

$$\hat{\beta} := (X'\{\widehat{Cov(y)}\}^{-1}X)^{-1}X'\{\widehat{Cov(y)}\}^{-1}y$$

für den unbekannten Parametervektor β.

LITERATURVERZEICHNIS

1. BALESTRA, P.: Best quadratic unbiased estimators of the variance-covariance matrix in normal regression. Journal of Econometrics 1, 17-28 (1973)

2. DRYGAS, H.: The Coordinate-Free Approach to Gauss-Markov Estimation
Berlin-Heidelberg-New York: Springer-Verlag 1970

3. DRYGAS, H.: The estimation of error-components: A note on a paper by Searle and Townsend.
Institut für Ökonometrie und Operations Research, Universität Bonn (1972)

4. HUBER, P.J.: The 1972 WALD memorial lectures
Robust regression: asymptotics, conjectures and monte carlo
The Annals of Statistics, Vol. 1, No.5, 799-821 (1973)

5. RAGHAVARAO, D.: Constructions and combinatorial problems in design of experiments. New York: Wiley 1971

6. RAO, C.R.: Minimum variance quadratic unbiased estimation of variance components.
Journal of Multivariate Analysis 1, 445-456 (1971)

7. RAO, C.R.: Estimation of variance and covariance components in linear models.
Journal of the American Statistical Association, 67, 112-115 (1972)

8. RAO, C.R.: Linear statistical inference and its applications. New York: Wiley 1973

9. SEARLE, S.R.: Linear models
New York: Wiley 1971

10. SEELY, J.: Linear spaces and unbiased estimation - application to the mixed linear model
The Annals of Math. Statistics, Vol.41, No.5, 1735-1748 (1970)

11. SEELY, J.: Quadratic subspaces and completeness.
The Annals of Math. Statistics, Vol. 42, No.2, 710-721 (1971)

12. SEELY, J.: Completeness for a family of multivariate normal distributions.
The Annals of Math. Statistics, Vol. 43, No.5 1644-1647 (1972)

13. SWAMY, P.A.V.B.: Statistical Inference in Random Coefficient Regression Models.
Berlin-Heidelberg-New York: Springer 1971

L e b e n s l a u f

Geburtsdatum	24. Januar 1942
und Geburtsort:	in Stuttgart
Eltern	Otto Urfer und Frau Maria, geb. Scheffel
Schulen und Studium:	1948 - 1953 Friedensschule Stuttgart
	1953 - 1962 Dillmann-Gymnasium Stuttgart
	1962 - 1968 Mathematik- und Physikstudium an der Universität Stuttgart
	1968 - 1969 Seminar für Studienreferendare Stuttgart
Berufstätigkeit:	1965 - 1968 Lehrer im Angestelltenverhältnis an der Waldschule in Stuttgart
	1968 - 1969 Studienreferendar und Assessor des Lehramtes am Goldberg-Gymnasium in Sindelfingen
	1970 - 1974 Wissenschaftlicher Assistent an der Abteilung für Angewandte Mathematik der Universität Hohenheim Seit 1.10.74 beim Institut für Mathematik an der Ruhr-Universität Bochum angestellt.

(Wolfgang Urfer)

Bochum, den 15.10.1975